1922–1923

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

ORMOND STONE
J. W. ALEXANDER
L. P. EISENHART
T. H. GRONWALL
OSWALD VEBLEN
J. H. M. WEDDERBURN

WITH THE COÖPERATION OF

A. A. BENNETT
H. BLUMBERG
G. A. PFEIFFER
J. K. WHITTEMORE

PUBLISHED BY THE

PRINCETON UNIVERSITY PRESS

SECOND SERIES, VOL. 24

LANCASTER, PA., AND PRINCETON, N. J.

1923

Lutcke & Wulff, Hamburg, Germany

PERIODICITIES IN THE THEORY OF PARTITIONS.

BY E. T. BELL.

I. Definitions and Fundamental Identities.

1. **Introduction.** If in the analysis of partitions we use theta functions, not merely theta constants as customary, we enter an extensive region of the theory of partitions which has not been worked hitherto, and which abounds in remarkable relations between Euler's $P(n)$, the total number of partitions of n, and six other denumerants introduced naturally by this extension. There is great flexibility in the method owing to the presence of continuous variables in the fundamental identities which are finite expressions containing denumerants. This is in sharp contrast to the usual analysis in which all relations between denumerants and functions of a continuous parameter involve infinite series or products.

The relations between denumerants fall into classes according to the periodicity of the coefficients. Two kinds of periodicity are encountered; the first is that familiar in analysis, the second is of the kind peculiar to the theory of numbers in which functions whose arguments differ by positive integral multiples of the period are not equal but are congruent with respect to certain moduli. Let us call the first of these ordinary periodicity and the second numerical. The type of numerical periodicity occurring in connection with partitions is that which appears in the integers defined by recurring series having scales of relation with integral coefficients the first of which is unity, or, what is equivalent, the algebraic numbers associated with the series are algebraic integers. This type was noticed by Lagrange, considered in detail for second order series by Lucas, and recently extended to the general case by Carmichael*.

In one class of relations some of the coefficients have ordinary periodicity and the rest are aperiodic; in the second only numerical periodicity appears; in a third some of the coefficients exhibit ordinary periodicity and the rest numerical, so that in a sense this class is doubly periodic; in a fourth class all of the coefficients are aperiodic. On account of their greater simplicity and interest we treat only linear relations; the non-linear are given by

* Quarterly Journal, vol. 48 (1920), pp. 343–372.

a parallel development starting from the addition theorems for the theta functions.

The entire subject is interconnected with that of Lucas' functions u_n, v_n and with cyclotomy, so that we have here a link between the division of the circle into equal parts and the theory of partitions of numbers. The division of the circle has already entered indirectly into partitions through Sylvester's theory of waves. The present application is distinct from this. It will be evident also that the subject is closely related to the transformation theory of the elliptic theta functions, particularly to that part which gives the formulas for the division of the argument; but we do not discuss this.

2. A-, E-, O-**partitions.** We consider those partitions of the positive integer n in each of which precisely r parts occur only once and no part appears more than twice, designating the number of such partitions by $A_r(n)$ if the parts may be odd or even, by $E_r(n)$ if all the parts are restricted to be even, and by $O_r(n)$ if all the parts are restricted to be odd; r is called the rank of each partition enumerated in $A_r(n)$, $E_r(n)$ or $O_r(n)$. The total number of partitions of n is denoted by $P(n)$, and the whole number of partitions of n into parts none of which appears more than twice is written $A(n)$ if the parts may be odd or even, $E(n)$ if all the parts are even, $O(n)$ if all the parts are odd.

Hence

$$A(n) = \sum A_r(n), \qquad E(n) = \sum E_r(n), \qquad O(n) = \sum O_r(n),$$

the summation extending to $r = 0, 1, 2, \cdots$. Clearly r can not exceed $1 + \frac{1}{2} n$.

By convention the value of each of the seven denumerants $A_r(n)$, $A(n), \cdots$, $O(n)$, $P(n)$ is 1 when $n = 0$.

From the definitions

$$E_r(2n+1) = O_{2r}(2n+1) = O_{2r+1}(2n) = 0,$$

and it is easily seen (at once from the generators in § 3, or from first principles) that $E_r(2n) = A_r(n)$. It is therefore unnecessary although sometimes convenient to retain both the A and E functions. We need also the following in which x is a parameter,

$$A'_n(x) = \sum x^r A_r(n), \qquad E'_n(x) = \sum x^r E_r(n), \qquad O'_n(x) = \sum x^r O_r(n),$$

the $\sum$ as before referring to $r = 0, 1, 2, \cdots$.

By convention $A'_0(x) = E'_0(x) = O'_0(x) = 1$, and from the definitions

$$A'_n(1) = A(n), \qquad E'_n(1) = E(n), \qquad O'_n(1) = O(n).$$

Of the seven functions $P(n), \cdots, O(n)$, only $A(n)$ and Euler's $P(n)$ seem to have been defined previously*.

$A(n)$ was introduced, discussed, and generalized by Glaisher†.

The remaining five can also be generalized in several ways immediately suggested by the forms of their generators, but as these extensions have no simple connection with the elliptic theta functions we leave them aside.

3. **Generators.** Indicate that the part α occurs precisely a times in a given partition by writing α_a. Then if $\alpha_a, \beta_b, \cdots, \gamma_c$ is a partition of n,

$$a\alpha + b\beta + \cdots + c\gamma = n,$$

and $\alpha, \beta, \cdots, \gamma$ are all different. If no part is to appear more than twice, each of $a, b, \cdots, c$ is one of the numbers 0, 1, 2. We may write $a \times n = n_a$, and hence

$$1 + xq^n + q^{2n} = q^{n_0} + xq^{n_1} + q^{n_2};$$

whence the generators

$$\prod (1 + xq^n + q^{2n}) = \sum q^n A'_n(x), \tag{1}$$

$$\prod (1 + xq^l + q^{2l}) = \sum q^n E'_n(x) = \sum q^{2n} A'_n(x), \tag{2}$$

$$\prod (1 + xq^m + q^{2m}) = \sum q^n O'_n(x), \tag{3}$$

in which the respective products extend to all integers $n > 0$, to all even integers $l > 0$, to all odd integers $m > 0$, and the summations to all integers $n > 0$. As obviously is permissible it is assumed that x, q both different from zero are such as to render the series and products absolutely convergent.

4. **Notation.** For brevity l, λ, m, μ, n, ν unless otherwise indicated represent integers of the following kinds: l, λ are even, m, μ odd, n, ν odd or even,

$$l > 0,\ \lambda \gtrless 0; \quad m > 0,\ \mu \gtrless 0; \quad n > 0,\ \nu \geqq 0.$$

The a^{th} triangular number, zero being counted the first, is written t_a: $t_a = \frac{1}{2} a(a-1)$, $a > 0$. The product sign $\prod$ extends to all integers except zero of the type indicated, the summation $\sum$ to all integers of the indicated type, zero included. Departures from these conventions will be specifically noted where they apply. *When the letter ν occurs under a $\sum$, the summation is with respect to ν only, and refers to $\nu = 0, 1, 2, \cdots$; similarly for a, the summation in this case extending to $a = 1, 2, 3, \cdots$.* From the nature of the functions all such sums are finite. Sums of denumerants whose arguments

* Cf. Dickson, History of the Theory of Numbers, vol. 2, chap. III; MacMahon, Combinatory Analysis, vol. 2.

† Messenger of Mathematics, vol. 12 (1883), p. 166.

are functions of λ, μ or ν are extended only to all those values of λ, μ or ν consistent with their definitions which render the arguments of the summands $\geqq 0$. The significance of any formula can be seen at once by referring to this paragraph.

5. **The fundamental identities.** In the usual notation of the elliptic theta functions

$$\vartheta_3(x) = q_0 \prod (1 + 2q^m \cos 2x + q^{2m}) = \sum q^{\nu^2} \cos 2\nu x,$$

$$q_0 = q_0(q) = \prod (1 - q^l), \qquad q_0^{-1} = \sum q^{2n} P(n),$$

and hence by (3),

$$q_0^{-1} \vartheta_3(\tfrac{1}{2} x) = \sum q^n O'_n (2 \cos x).$$

Recalling that $O_r(n) = 0$ when r, n are not both even or both odd we equate coefficients of like powers of q in the last identity distinguishing the cases $n = l$, $n = m$:

$$(4) \qquad \sum 2^{2r} O_{2r}(l) \cos^{2r} x = \sum P\left(\frac{l - \lambda^2}{2}\right) \cos \lambda x.$$

$$(5) \qquad \sum 2^{2r+1} O_{2r+1}(m) \cos^{2r+1} x = \sum P\left(\frac{m - \mu^2}{2}\right) \cos \mu x.$$

In the same way from

$$\vartheta_2(x) = 2 q_0 q^{\frac{1}{4}} \cos x \prod (1 + 2q^l \cos 2x + q^{2l}) = \sum q^{\frac{\mu^2}{4}} \cos \mu x.$$

we find, on dividing through by q_0 and using (2),

$$(6) \qquad 2 \cos x \sum 2^r E_r \left(\frac{m-1}{4}\right) \cos^r 2x = \sum P\left(\frac{m - \mu^2}{8}\right) \cos \mu x$$

when and only when $m \equiv 1$ mod. 8. These three can be written also

$$(4.1) \qquad \sum 2^{2r} O_{2r}(2n) \cos^{2r} x = P(n) + 2 \sum P(n - 2a^2) \cos 2ax.$$

$$(5.1) \qquad \sum 2^{2r+1} O_{2r+1}(2n+1) \cos^{2r+1} x = 2 \sum P(n - 4t_a) \cos (2a-1) x.$$

$$(6.1) \qquad \cos x \sum 2^r A_r(n) \cos^r 2x = \sum P(n - t_a) \cos (2a-1) x.$$

where in the last we have replaced $E_r(2n)$ by $A_r(n)$. Multiply (6.1) throughout by $\cos x$ and replace x by $\tfrac{1}{2}x$.

$$(6.2) \qquad 2A_0(n) + \sum 2^a [A_{a-1}(n) + 2 A_a(n)] \cos^a x$$
$$= 2P(n) + 2 \sum [P(n - t_a) + P(n - t_{a+1})] \cos a x.$$

In (4.1) — (6.1) change x into $x + \frac{1}{2}\pi$,

$$(4.11)\quad \sum 2^{2r}\theta_{2r}(2n)\sin^{2r}x = P(n) + 2\sum(-1)^a P(n-2a^2)\cos 2ax.$$

$$(5.11)\quad \sum 2^{2r+1}\theta_{2r+1}(2n+1)\sin^{2r+1}x$$
$$= -2\sum(-1)^a P(n-4t_a)\sin(2a-1)x.$$

$$(6.11)\quad \sin x\sum(-1)^r 2^r A_r(n)\cos^r 2x$$
$$= -\sum(-1)^a P(n-t_a)\sin(2a-1)x;$$

and in (6.2) change x into $\pi - x$,

$$(6.21)\quad \begin{aligned} &2A_0(n) + \sum(-1)^a 2^a[A_{a-1}(n) + 2A_a(n)]\cos^a x \\ &= 2P(n) + 2\sum(-1)^a[P(n-t_a) + P(n-t_{a+1})]\cos ax. \end{aligned}$$

6. **Introduction of Lucas' functions***. It follows from De Moivre's theorem that if p, q are real and x is the real or imaginary angle such that

$$(7)\qquad 2\sqrt{q}\cos x = p, \qquad 2\sqrt{q}\sin x = \sqrt{4q-p^2},\ 4q \neq p^2\dagger,$$

then for all values > 0 of n,

$$(8)\qquad 2q^{\frac{n}{2}}\cos nx = v_n, \qquad 2q^{\frac{n}{2}}\sin nx = u_n\sqrt{4q-p^2},$$

where u_n, v_n is that pair of linearly independent solutions of the difference equation

$$(9)\qquad y_{n+2} = py_{n+1} - qy_n$$

which is defined by the initial values

$$(10)\qquad u_0 = 0,\ u_1 = 1,\ v_0 = 2,\ v_1 = p.$$

When p, q are integers $\gtrless 0$, u_n, v_n are the singly periodic numerical functions of Lucas, the implied periodicity in this definition being numerical

* E. Lucas, (1) American Journal of Mathematics, vol. 1 (1878) pp. 184–238, 289–321; (2) Theorie des Nombres, chap. 18, cf. also Bachmann, Niedere Zahlentheorie, vol. 2, chap. 3; Dickson, loc. cit. vol. 1, chap. 17, also papers of Cipolla, ibid. p. 218 for properties relating to the periodicity. Note that in (2) Lucas abandons the restriction that p, q be relatively prime in the definition of u_n, v_n, while Bachmann retains it. The implications of this distinction do not concern us here; cf. Carmichael, loc. cit. § 2.

† If $4q = p^2$ the results are developed otherwise, cf. § 18.

as defined in § 1 and independent of the period $2\pi/n$ of the circular functions in (8). For assigned values of p, q the moduli with respect to which this periodicity obtains are determinate, as also are the sequences u_n, v_n $(n = 0, 1, \ldots)$. The particular sequences defined by $p = a$, $q = b$ (a, b constants) are called the u, v sequences for (a, b). To eight of the most interesting Lucas gave the following names: $(1, -1)$, Fibonacci; $(2, -1)$, Pell; $(2, -3)$ conjugate Pell; $(3, 2)$ Fermat. As will be seen later these cases play an important in the theory of partitions. The Fibonacci u-sequence $0, 1, 1, 2, 3, 5, \ldots$ is usually called Fibonacci's numbers, although Bachmann gives this name to the v-sequence $2, 1, 3, 4, 7, 11, \ldots$, calling the first Lamé's. He also names Pell's sequences after Dupré. When p, q are integers we shall call u_n, v_n $(n = 0, 1, 2, \ldots)$ the Lucas' u, v sequences for (p, q).

In what immediately follows p, q are arbitrary real quantities. Write $\delta = \sqrt{p^2 - 4q}$, so that

$$\delta \frac{\partial \delta}{\partial p} = p, \qquad \delta \frac{\partial \delta}{\partial q} = -2,$$

and hence from (7), (8):

$$(11) \qquad \frac{\partial v_n}{\partial p} = n u_n, \qquad \frac{\partial v_n}{\partial q} = -n u_{n-1},$$

$$(12) \qquad \delta \frac{\partial (\delta u_n)}{\partial p} = n v_n, \qquad \delta \frac{\partial (\delta u_n)}{\partial q} = -n v_{n-1},$$

$$(13) \qquad \int u_n \, dp = 1/n \cdot v_n, \qquad \int u_n \, dq = -1/(n+1) \cdot v_{n+1},$$

$$(14) \qquad 1/\delta \cdot \int v_n/\delta \, dp = 1/n \cdot u_n, \qquad 1/\delta \cdot \int v_n/\delta \, dq = -1/(n+1) \cdot u_{n+1},$$

and the explicit values of u_n, v_n are given by

$$(15) \qquad 2^n \delta u_n = (p+\delta)^n - (p-\delta)^n, \qquad 2^n v_n = (p+\delta)^n + (p-\delta)^n,$$

these being a pair of fundamental solutions of (9).

7. Ordinary periodicity. This occurs in connection with partitions as follows. Let $\alpha(n)$ be a function which takes a single integral value for each value ≥ 0 of the integer n. Then if there exists a finite integer $n' > 0$ such that for k an arbitrary integer ≥ 0, $\alpha(n + kn') = \alpha(n)$ $(n = 0, 1, 2, \cdots)$, n' is a period of $\alpha(n)$; and if m is the least period of $\alpha(n)$ the values of $\alpha(n)$

recur in cycles of m and in no smaller cycle. Let us call this least cycle of values the recurrence of $\alpha(n)$ and indicate it thus,

$$\begin{aligned} n &= 1, 2, 3, \cdots, m-1, 0 \bmod. m, \\ \alpha(n) &= \alpha(1), \alpha(2), \alpha(3), \cdots, \alpha(m-1), \alpha(0), \end{aligned}$$

giving in any specific case the actual values of $\alpha(1), \alpha(2), \cdots, \alpha(0)$. This kind of periodicity appears in sums of the forms

$$\sum \alpha(a)\, P(n - 2a^2), \qquad \sum \alpha(a)\, P(n - c t_a), \qquad c = 1, 4.$$

Such sums are not periodic functions of m or of n; the periodicity is only in the cyclic recurrence of the coefficients. When m is the period of $\alpha(n)$ it is convenient to say that each of these sums is of period m.

To see more clearly the nature of relations involving $\alpha(n)$ functions consider the sum of period 5,

$$\begin{aligned} -P(n-1) + 2P(n-3) - P(n-6) - P(n-21) \\ + 2P(n-28) - P(n-36) - P(n-66) + \cdots, \end{aligned}$$

which will be shown equal to

$$-A_1(n) + A_2(n) - 2A_3(n) + 3A_4(n) - 5A_5(n) + 8A_6(n) - 13A_7(n) + \cdots,$$

the coefficients 1, 1, 2, 3, 5, 8, $\cdots$ being Fibonacci's u. The arguments of the P's are of the form $n - t_a$ and decrease; the argument n of the A's is constant, and the rank variable. Thus the periodic (ordinary) sum is reduced to a linear function of A-partitions of n alone, the coefficients in which are numerically periodic. The relation can be written

$$\sum \alpha(a)\, P(n - t_a) = \sum (-1)^r u_r A_r(n),$$

where u_r is the r^{th} Fibonacci number, and the recurrence of $\alpha(n)$ is

$$\begin{aligned} a &= 1,\ 2,\ 3,\ 4,\ 0 \bmod. 5, \\ \alpha(a) &= 0, -1,\ 2, -1,\ 0. \end{aligned}$$

Fibonacci's u_{r+2} has already appeared in the theory of partitions, MacMahon (loc. cit. p. 46) having introduced it from other considerations to enumerate the whole number of partitions of all numbers having no sequences, no repetitions and no part greater than r. It may be possible therefore to interpret

the foregoing sum of A-functions in terms of partitions of n of one particular kind. The like applies to sums involving binomial coefficients (§ 18), for the latter also have assigned meanings as denumerants by MacMahon (ibid. p. 48).

II. Relations in which Lucas' Numerically Periodic Coefficients Occur.

8. **General relations.** The u_n, v_n in (I)-(VIII) are any solutions of (9) determined by the initial conditions (10). Substitute in (4.1)-(6.21) the values of $\cos x$, $\cos ax$ from (7), (8); to get (IV), (VIII) add and subtract the pair thus found from (6.2), (6.21); write out the others directly from the substitution:

(I) $$\sum (p^2/q)^r O_{2r}(2n) = P(n) + \sum \frac{v_{2a}}{q^a} P(n-2a^2).$$

(II) $$p\sum (p^2/q)^r O_{2r+1}(2n+1) = q\sum \frac{v_{2a-1}}{q^a} P(n-4t_a).$$

(III) $$p\sum (p^2/q-2)^r A_r(n) = q\sum \frac{v_{2a-1}}{q^a} P(n-t_a).$$

(IV) $$2A_0(n) + \sum (p^2/q)^a [A_{2a-1}(n) + 2A_{2a}(n)]$$
$$= 2P(n) + \sum \frac{v_{2a}}{q^a} [P(n-t_{2a}) + P(n-t_{2a+1})].$$

(V) $$\sum (4-p^2/q)^r O_{2r}(2n) = P(n) + \sum (-1)^a \frac{v_{2a}}{q^a} P(n-2a^2).$$

(VI) $$\sum (4-p^2/q)^r O_{2r+1}(2n+1) = -q\sum (-1)^a \frac{u_{2a-1}}{q^a} P(n-4t_a).$$

(VII) $$\sum (2-p^2/q)^r A_r(n) = -q\sum (-1)^a \frac{u_{2a-1}}{q^a} P(n-t_a).$$

(VIII) $$\sum (p^2/q)^a [A_{2a-2}(n) + 2A_{2a-1}(n)]$$
$$= \sum \frac{v_{2a-1}}{q^a} [P(n-t_{2a-1}) + P(n-t_{2a})].$$

Each of these generates an infinite chain of results of the same general type upon repeated application of (11)-(14) in any order. At the n^{th} step there are possible $8 \times 7^{n-1}$ different relations deduced from any one. For at each step may be applied any one of the eight processes in (11)-(14) except the inverse of that last used. Thus the p-derivative of (I) is

$$\sum r(p^2/q)^r O_{2r}(2n) = p\sum \frac{a\,u_{2a}}{q^a} P(n-2a^2);$$

multiply this by δ, take again the p-derivative and multiply by δ,

$$\sum[r^2(p^2/q)^r + 2r(2r-1)(p^2/q)^{r-1}]O_{2r}(2n) = \sum \frac{a^2 v_{2a}}{q^a} P(n - 2a^2),$$

and so on.

9. **Relations with Lucas' Functions.** Thus far p, q are arbitrary real quantities. When p, q are integers, the u, v in all of the above relations are Lucas' functions. For the specially named sequences in § 6 the results are particularly simple and interesting, but we need not write them out.

In the case of arbitrary p, q it is clear that in any of the relations coefficients of p^a, q^b can be equated after u, v have been replaced by their explicit forms (15) and expanded. The result is a chain of aperiodic relations. Some of the more important of these are obtained otherwise in § 18.

III. Ordinary Periodicity of Odd Prime Period p.

10. **Preliminary reductions of cyclotomic sums.** This class of relations depends upon the $e = \frac{1}{2}(p-1)$ binomial periods

$$\eta'_k = r^{g^k} + r^{-g^k} = 2\cos g^k \theta_p \, (0 \overline{\overline{<}} k < e)$$

appertaining to the p-section of the circle*, p an odd prime, $\theta_p = 2\pi/p$, $g =$ a primitive root of p. In what follows it is immaterial which g be selected, but the choice once made is to be permanent, and likewise for the period η chosen presently. As j runs through all integers > 0, η'_j takes only the e distinct values $\eta'_k (0 \overline{\overline{<}} k < e)$.

The values of $2\cos a\theta_p (0 < a \overline{\overline{<}} e)$ are all different, and if a' is the least positive index of a to the base g,

$$\eta'_a = 2\cos a\theta_p = 2\cos(p-a)\theta_p. \qquad (0 < a \overline{\overline{<}} e)$$

It follows that $2\cos a\theta_p (0 < a \overline{\overline{<}} e)$ is a permutation of $\eta'_k (0 \overline{\overline{<}} k < e)$. Write $\eta'_a = \eta_a$, so that the e binomial periods are

$$\eta_a = 2\cos a\theta_p = 2\cos(p-a)\theta_p. \qquad (0 < a \overline{\overline{<}} e)$$

By convention η_0 (not one of the periods) has the value 2. Reduced forms of four sums will be required.

Let η be a particular one of the binomial periods and $\lambda_r(n) (r = 0, 1, 2, \cdots)$ single valued functions real for all values $\geqq 0$ of the integer n.

* Cf. Mathews, Theory of Numbers, chap. 7; Bachmann, Kreistheilung, chap. 6. There is no corresponding class based on the f-nomial periods when $f > 2$.

Then

$$\lambda(n) = \sum \eta^r \lambda_r(n) \tag{16}$$

is uniquely reducible to

$$\lambda(n) = \lambda_0(n) + L_1(n)\eta_1 + L_2(n)\eta_2 + \cdots + L_e(n)\eta_e, \tag{17}$$

in which $L_s(n)(0 < s < e)$ are real. The appropriate reduction of $L_s(n)$ being rather long we first state the result for clearness. Since η is a *particular* period, one and only one set of $e(e-1)$ integers c_{jk} is defined by the identities

$$\eta^j = c_{j1}\eta_1 + c_{j2}\eta_2 + \cdots + c_{je}\eta_e \qquad (0 < j < e) \tag{18}$$

obtained by expressing each of $\eta, \eta^2, \cdots, \eta^{e-1}$ as a linear homogeneous function of all the periods; moreover the e periods are the roots of an irreducible equation

$$\eta^e = a_{e-1}\eta^{e-1} + a_{e-2}\eta^{e-2} + \cdots + a_1\eta + a_0 \tag{19}$$

with integral coefficients. The e^2 integers

$$c_{js},\ a_k (0 < j < e,\ 0 < s \overline{\overline{<}} e,\ 0 \leqq k < e), \tag{20}$$

together with the recurrence relation of order e

$$f(n+e) = a_{e-1}f(n+e-1) + a_{e-2}f(n+e-2) + \cdots + a_0 f(n), \tag{21}$$

whose associated algebraic equation is the equation (19) of the periods, determine $L_s(n)$ for all values > 0 of the integer n. For it will be shown that

$$L_s(n) = \sum \lambda_a(n) u_s(a), \qquad (0 < s < e) \tag{22}$$

the summation by § 4 being with respect to $a = 1, 2, 3, \cdots$, where $u_s(n)(n \geqq 0)$ is that solution of (21) which is given by the $e-1$ initial values

$$u_s(j) = c_{js} (0 < j < e). \tag{23}$$

To prove this reduce (16) by (19) to

$$\lambda(n) = \lambda_0(n) + A_1(n)\eta + A_2(n)\eta^2 + \cdots + A_{e-1}(n)\eta^{e-1}, \tag{24}$$

in which the A's and λ_0 are independent of η, and write

$$\eta^n = q_{e-1}(n)\eta^{e-1} + q_{e-2}(n)\eta^{e-2} + \cdots + q_0(n). \tag{25}$$

the q's being integers. Multiply (25) throughout by η, reduce the right of this result by the right of (19), observe that by (25) the left of the identity so obtained is

$$q_{e-1}(n+1)\eta^{e-1} + q_{e-2}(n+1)\eta^{e-2} + \cdots + q_0(n+1),$$

and equate coefficients of like powers of η, getting

$$(26)\quad q_0(n+1) = a_0 q_{e-1}(n), \quad q_j(n+1) = a_j q_{e-1}(n) + q_{j-1}(n). \quad (0<j<e)$$

From (26) it follows that

$$(27)\qquad q_0(n), \quad q_1(n), \cdots, \quad q_{e-1}(n)$$

is that set of e linearly independent solutions of (21) which is defined by the system of initial values

$$(28)\quad q_j(j) = 1\ (0 \leqq j < e), \quad q_j(k) = 0\ (0 \leqq j < e, \quad k < e, \quad j \neq k);$$

that is, (27) is the fundamental set of solutions of (21). Substitute in (16) the values (25) of η, η^2, $\cdots$ and compare with (24):

$$(29)\qquad A_j(n) = \sum \lambda_a(n)\, q_j(a), \qquad (0<j<e)$$

the $\sum$ by § 4 referring to $a = 1, 2, 3, \cdots$. In (24) substitute for $A_j(n)$ from (29), η^j ($j = 1, 2, \cdots, e-1$) from (18), and compare with (17). The coefficient of $\lambda_a(n)$ in $L_s(n)$, that is $u_s(a)$, is

$$(30)\qquad u_s(a) = \sum_{j=1}^{e-1} c_{js}\, q_j(a);$$

whence by comparing (30), (28) we have the stated initial conditions (23), which completes the verification of (22).

With this is a complementary reduction obtained from the preceding by changing the sign of each period. The constants (20) are as before, and we find immediately that

$$(16.1)\qquad \mu(n) = \sum (-1)^s \eta^s \mu_s(n) = \mu_0(n) + \sum (-1)^a \eta^a \mu_a(n)$$

can be reduced to

$$(17.1)\qquad \mu(n) = \mu_0(n) + M_1(n)\,\eta_1 + M_2(n)\,\eta_2 + \cdots + M_e(n)\,\eta_e,$$

where now

$$(22.1)\qquad M_s(n) = \sum \mu_a(n)\, v_s(a), \qquad (0<s \leqq e)$$

and $c_s(n)$ is that solution of

$$h(n+e) = \sum_{j=1}^{e} (-1)^j a_{e-j} h(n+e-j), \tag{21.1}$$

which is defined by the $e-1$ initial values

$$c_s(j) = (-1)^{j-1} c_{js}; \qquad (0<j<e) \tag{23.1}$$

the formulas being numbered correspondingly to those in the first reduction.

The next type of sum, depending on the values of $2\cos a\theta_p$ $(a = 1, 2, 3, \cdots)$, introduces the $p+1$ elementary periodic functions upon which the relations of period p are based. Write $[x]$ for the greatest integer not greater than x, hence $e = [\frac{1}{2}p]$, and define $e+1$ periodic functions $\alpha_k(a)$ $(0 \overline{\gtrless} k \overline{\gtrless} e)$ by

$$\begin{aligned} \alpha_{2j}(a) &= 1, \quad a \equiv \pm j \bmod. p, \quad 0 < j < [\tfrac{1}{4}p]; \\ \alpha_{p-2j}(a) &= 1, \quad a \equiv \pm j \bmod. p, \quad [\tfrac{1}{4}p] < j \overline{\gtrless} [\tfrac{1}{2}p]; \\ \alpha_0(a) &= 1, \quad a \equiv 0 \bmod. p, \end{aligned} \tag{31}$$

the values of $\alpha_k(a)$ in all cases except these being zero. Then it is easy to verify that

$$2\cos 2a\theta_p = \alpha_k(a)\,\eta_k \qquad (0 \overline{\gtrless} k \overline{\gtrless} e) \tag{32}$$

for all values >0 of the integer a. We recall that η_0 has the conventional value 2. In the same way $e+1$ periodic functions $\beta_k(a)$ $(0 \overline{\gtrless} k \overline{\gtrless} e)$ are defined by

$$\begin{aligned} \beta_{2j-1}(a) &= 1, \quad a \equiv j, 1-j \bmod. p, \quad 0 < j \overline{\gtrless} [\tfrac{1}{4}(p+1)]; \\ \beta_{p-2j+1}(a) &= 1, \quad a \equiv j, 1-j \bmod. p, \quad [\tfrac{1}{4}(p+1)] < j < [\tfrac{1}{2}p]; \\ \beta_0(a) &= 1, \quad a \equiv \tfrac{1}{2}(p+1) \bmod. p, \end{aligned} \tag{33}$$

the values of $\beta_k(a)$ in all other cases being zero; and for all values of the integer $a>0$,

$$2\cos(2a-1)\theta_p = \beta_k(a)\,\eta_k. \qquad (0 \overline{\gtrless} k \overline{\gtrless} e) \tag{34}$$

Let ψ_0, χ_0 be real quantities, $\psi(a)$, $\chi(a)$ functions which are single valued and real for each integer $a>0$. Then by (32) the sum

$$\psi_0 + 2\sum\psi(a)\cos 2a\theta_p \tag{35}$$

is reduced to

$$\Psi_0 + \Psi_1\eta_1 + \Psi_2\eta_2 + \cdots + \Psi_e\eta_e, \tag{36}$$

in which the Ψ's are real, and

$$(37)\qquad \Psi_0 = \psi_0 + 2\sum \alpha_0(a)\,\psi(a), \qquad \Psi_j(a) = \sum \alpha_j(a)\,\psi(a). \qquad (0 < j \leqq e)$$

Similarly by (34),

$$(38)\qquad \chi_0 + 2\sum \chi_a \cos(2a-1)\,\theta_p = X_0 + \sum_{j=1}^{e} X_j\,\eta_j,$$

$$(39)\qquad X_0 = \chi_0 + 2\sum \beta_0(a)\,\chi(a). \qquad X_j = \sum \beta_j(a)\,\chi(a). \qquad (0 < j \leqq e)$$

11. **Systems of e relations each of odd prime ordinary period p.** Take $\eta = \eta_1$ in (18), *thus fixing the integers c_{js} in* (20). These values are understood in the rest of this section. The a_k in (20) are the same for all choices of η. In each of (IX)-(XII) below j is defined by $0 < j \leqq e$, so that each equality represents $e = \frac{1}{2}(p-1)$ relations in which the functions α', β', γ' are ordinary periodic of odd prime period p, and by Carmichael's work cited in § 1, the functions u, v have numerical periodicity.

In (4.1) put $x = \theta_p$, hence $2\cos x = \eta_1$, and reduce the result by (17), (36) and $\sum_{j=1}^{e} \eta_j = -1$ to

$$\sum_{j=1}^{e}\left[-O_0(2n) + \sum u_j(2a)\,O_{2a}(2n)\right]\eta_j$$
$$= \sum_{j=1}^{e}\left[-P(n) + \sum\{\alpha_j(a) - 2\alpha_0(a)\}\,P(n-2a^2)\right]\eta_j.$$

Write $\alpha_j(a) - 2\alpha_0(a) = \alpha'_j(a)$ and in the preceding equate coefficients of η_j:

$$(\mathrm{IX})\qquad O_0(2n) - \sum u_j(2a)\,O_{2a}(2n) = P(n) - \sum \alpha'_j(a)\,P(n-2a^2),$$

in which, from (31), $\alpha'_k(a) = 0$ $(0 < k < e)$ except in the cases

$$(40)\qquad \begin{array}{llll} a \equiv \pm j \bmod. p, & 0 < j < [\frac{1}{4}p], & \alpha'_{2j}(a) = 1, \\ a \equiv \pm j \bmod. p, & [\frac{1}{4}p] < j \leqq [\frac{1}{2}p], & \alpha'_{p-2j}(a) = 1, \\ a \equiv 0 \bmod. p, & 0 < j \leqq [\frac{1}{2}p], & \alpha'_j(a) = -2. \end{array}$$

The u-coefficients are integers determined by (21), (23).

For the same u's, (5.1) compared with (37)-(39) yields the system

$$(\mathrm{X})\qquad \sum u_j(2a-1)\,O_{2a-1}(2n+1) = \sum \beta'_j(a)\,P(n-4t_a);$$

$\beta_k(a) = 0 (0 < k < e)$ except in the cases

$$
(41)\quad
\begin{array}{lll}
a \equiv j,\ 1-j \bmod. p, & 0 < j \leqq [\frac{1}{4}(p+1)], & \beta'_{2j-1}(a) = 1. \\
a \equiv j,\ 1-j \bmod. p, & [\frac{1}{4}(p+1)] < j \leqq [\frac{1}{2}p], & \beta'_{p-2j+1}(a) = 1. \\
a \equiv \frac{1}{2}(p+1) \bmod. p, & 0 < j \leqq [\frac{1}{2}p], & \beta'_j(a) = -2.
\end{array}
$$

In (6.2) put $x = \theta_p$ and get for the same α's the system

$$
\begin{aligned}
(\mathrm{XI})\quad 2A_0(n) - \sum \alpha_j(a)[A_{a-1}(n) + 2A_a(n)] \\
= 2P(n) - \sum \gamma'_j(a)[P(n-t_a) + P(n-t_{a+1})];
\end{aligned}
$$

$\gamma'_k(a) = 0 (0 < k < e)$ except in the cases

$$
(42)\quad
\begin{array}{lll}
a \equiv \pm j \bmod. p, & 0 < j \leqq [\frac{1}{2}p], & \gamma'_j(a) = 1; \\
a \equiv 0 \bmod. p, & 0 < j < [\frac{1}{2}p], & \gamma'_j(a) = -2.
\end{array}
$$

Comparing (6.21) with (16.1) we find for the same γ' the system

$$
\begin{aligned}
(\mathrm{XII})\quad 2A_0(n) + \sum c_j(a)[A_{a-1}(n) + 2A_a(n)] \\
= 2P(n) - \sum (-1)^a \gamma'_j(a)[P(n-t_a) + P(n-t_{a+1})],
\end{aligned}
$$

where the c's are integers determined by (21.1), (23.1).

13. The recurrence (§ 7) of each of the periodic functions $\alpha'_j(a)$, $\beta'_j(a)$, $\gamma'_j(a)$ consists of p terms one of which is -2 and of the remaining $p-1$ terms two are units $(+1)$ and the rest zeros. In the recurrences of $\alpha'_j(a)$, $\gamma'_j(a)$ the terms, the last which is -2 being omitted, read the same backwards and forwards; the recurrence of $\beta'_j(a)$ is symmetrical about its middle term -2. Each of the systems (IX)-(XII) determines a relation in which the recurrence for the periodic function consists of $[\frac{1}{2}p]$ arbitrary constants k_j and retains its symmetry as above. In these the coefficients a or c on the left are again given by the recurring series (21) or (21.1) but with sets of $e-1$ initial values depending linearly on the k_j. If the constants k_j are integers the a, c are, by Carmichael's work as before, numerically periodic. It will be sufficient to write out the relation of this type deduced from (IX): the remaining three are found in the same way.

In (IX) replace j by $2j$, multiply by k_{2j} and (glancing at (40)) put in turn for j each of the integers $0 < j \leqq [\frac{1}{4}p]$, in (IX) replace j by $p-2j$ and let j range $[\frac{1}{4}p] < j < [\frac{1}{2}p]$; add all the relations thus obtained. Write $k_0 = 2(k_1 + k_2 + \cdots + k_e)$. Then if $\alpha(a)$ is the periodic function whose recurrence is

$$
(43)\quad
\begin{array}{l}
a \equiv 1, 2, 3, \cdots, p-3, p-2, p-1, 0 \bmod. p, \\
\alpha(a) = k_1, k_2, k_3, \cdots, k_3, k_2, k_1, -2k_0,
\end{array}
$$

and $u(r)$ $(r = 1, 2, 3, \cdots)$ is that solution of (21) which is determined by the initial values

$$u(r) = \sum_{j=1}^{e} k_j c_{rj} \quad (0 < r < e), \tag{44}$$

we have for the sum just constructed

$$k_0 O_0(2n) - \sum u(2a) O_{2a}(2n) = k_0 P(n) - \sum u(a) P(n - 2a^2). \tag{XIII}$$

IV. Relations Involving the Numbers of Fibonacci, Pell, Fermat and Lucas.

14. **Dependence upon quadratic integers.** In section II the Lucas numbers appear as coefficients of the P functions; in the class to be considered next these numbers are coefficients of the A, O functions and the coefficients of the P's are ordinary-periodic. Let d, p_0, q_0, p_1, q_1 be rational integers, of which $d > 1$ has no square divisor > 1; p_0, q_0 are arbitrary, and $p_1 \equiv q_1$ mod. 2. All integers ζ_0, ζ_1 of the corpus $K(\sqrt{d})$ fall into the classes*

$$d \equiv 2, 3 \text{ mod. } 4, \qquad \zeta_0 = p_0 + q_0\sqrt{d}; \tag{45}$$

$$d \equiv 1 \text{ mod. } 4, \qquad \zeta_1 = \frac{p_1 + q_1\sqrt{d}}{2}. \tag{46}$$

Let u_n, v_n $(n = 0, 1, \cdots)$ be the Lucas sequences (§ 6) for

$$(p, q) = (2p_0, p_0^2 - d\,q_0^2); \tag{45.1}$$

then it follows that

$$2\zeta_0^n = v_n + q_0\sqrt{d}\,u_n; \tag{45.2}$$

and if u_n, v_n $(n = 0, 1, \cdots)$ are the Lucas sequences for

$$(p, q) = \left(p_1, \frac{p_1^2 - d\,q_1^2}{2}\right), \tag{46.1}$$

then

$$2\zeta_1^n = v_n + q_1\sqrt{d}\,u_n. \tag{46.2}$$

Considering first the primitive intergers of $K(\sqrt{d})$ we take $q_0, q_1 \neq 0$, and find in what cases $a \cos^c bq_k$, $a \sin^c bq_k$ are quadratic integers, a, b, c, k being rational integers > 0 and $q_k = \pi/k$. If q_0 or $q_1 < 0$, we take $-\zeta_0$ or $-\zeta_1$

* Cf. Bachmann, Allgemeine Arithmetik, chap. 5; Reid, Algebraic Numbers, chap. 10.

instead of ζ_0 or ζ_1 in (45.2), (46.2); and if $p_0 = d$ or $p_1 = d$ we take $\zeta_0/\sqrt{d}$, $\zeta_1/\sqrt{d}$ after having made where necessary the foregoing change of sign. These changes give the essentially simplest Lucas sequences; otherwise we get finally a different but equivalent set of relations between denumerants*.

We find the following eight cases, classified according to the sequences involved.

$(p, q) = (1, -1)$ *Fibonacci u, v*:

$$
(47)\quad
\begin{aligned}
2^{2n-1} \cos^{2n} \varphi_{10} &= 2^{2n+1} \sin^{2n} 2\varphi_5 = (\sqrt{5})^n (v_n + \sqrt{5}\,u_n);\\
2^{n-1} \cos^{n} \varphi_5 &= 2^{n-1} \sin^{n} 3\varphi_{10} = v_n + \sqrt{5}\,u_n;\\
2^{2n-1} \cos^{2n} 3\varphi_{10} &= 2^{2n+1} \sin^{2n} \varphi_5 = (-\sqrt{5})^n (v_n - \sqrt{5}\,u_n);\\
2^{n+1} \cos^{n} 2\varphi_5 &= 2^{n+1} \sin^{n} \varphi_{10} = (-1)^n (v_n - \sqrt{5}\,u_n).
\end{aligned}
$$

$(p, q) = (2, -1)$: *Pell u, v*:

$$
(48)\quad
\begin{aligned}
2^{2n-1} \cos^{2n} \varphi_8 &= 2^{2n+1} \sin^{2n} 3\varphi_8 = (\sqrt{2})^n (v_n + \sqrt{2}\,u_n);\\
2^{2n-1} \cos^{2n} 3\varphi_8 &= 2^{2n-1} \sin^{2n} \varphi_8 = (-\sqrt{2})^n (v_n - \sqrt{2}\,u_n).
\end{aligned}
$$

$(p, q) = (4, 1)$: *Lucas u, v*:

$$
(49)\quad
\begin{aligned}
2^{2n+1} \cos^{2n} \varphi_{12} &= 2^{2n-1} \sin^{2n} 5\varphi_{12} = v_n + \sqrt{3}\,u_n;\\
2^{2n-1} \cos^{2n} 5\varphi_{12} &= 2^{2n+1} \sin^{2n} \varphi_{12} = v_n - \sqrt{3}\,u_n.
\end{aligned}
$$

The last case, $(p, q) = (4, 1)$, has been named after Lucas for convenience. These three cases correspond to the division of the circle into 10 or 20, 16, 24 parts; and for the divisions into 2, 4, 6, 8, 12 parts we have $\cos^n \varphi_1 = (-1)^n$, $\sin^n \varphi_2 = 1$:

$$
(50)\quad
\begin{aligned}
2^n \cos^n \varphi_3 &= 2^n \sin^n \varphi_6 = 1;\\
2^n \cos^n \varphi_4 &= 2^n \sin^n \varphi_4 = (\sqrt{2})^n;\\
2^n \cos^n \varphi_6 &= 2^n \sin^n \varphi_3 = (\sqrt{3})^n.
\end{aligned}
$$

In the Fermat case $(p, q) = (3, 2)$ $u_n = 2^n - 1$, $v_n = 2^n + 1$;

$(p, q) = (3, 2)$, *Fermat u, v*:

$$
(51)\quad
\begin{aligned}
2^n [\cos^n \varphi_1 + (-1)^n \cos^n \varphi_3] &= (-1)^n u_n, \ (-1)^n v_n;\\
2^n (\sin^n \varphi_2 + \sin^n \varphi_6) &= u_n, v_n.
\end{aligned}
$$

* In checking the next list the table in Hobson's Trigonometry (Cambridge, 1897), p. 72 will be found useful.

Recapitulating: when $k = 1, 2, 3, 4, 5, 6, 8, 10, 12$, we can choose rational integers $a, b, c, > 0$ so that $a \cos^c b\varphi_k$, $a \sin^c b\varphi_k$ are quadratic integers, and the values of these functions then are given by (47) — (51).

The generalization to the case of algebraic integers of degree > 2 is sufficiently obvious. We do not consider this now, as clearly the associated u, v are not Lucas' sequences.

15. **Periodic functions.** To have readily accessible the data for a complete discussion of the relations in which the associated algebraic integers are of the second degree we assemble here all the ordinary periodic functions occurring first. From these, which form a sort of fundamental set, all remaining periodic functions belonging to this case are constructed as in the derivations in § 16. To save space we adopt an abridged notation. Write the elements in the recurrence of a periodic function $\alpha(a)$ in the order in which they occur, beginning with the element which gives the value of the function when the argument is congruent to 1 for the given modulus m. If any element, $-a$, is negative, write the minus sign above it, $\bar{a}$; and if an element a (or b) is repeated j times in succession, write it a_j (or b_j), and similarly for repeated sequences of elements; thus $(ab)_j$ means $ababab\cdots$ to j terms ab. Finally enclose the result in $(\ )_m$ the suffix indicating the modulus, and write $\alpha(a) = (*)_m$, the $*$ denoting the elements written as above described. For example the recurrence

$$
\begin{array}{rcl}
a & = & 1, 2, 3, 4, \quad 5, \quad 6, 7, 8, 9, 0 \text{ mod. } 10, \\
\alpha(a) & = & 1, 0, 0, 0, -1, -1, 0, 0, 0, 1
\end{array}
$$

is abridged to $\alpha(a) = (1\, 0_3\, 1_2\, 0_3\, 1)_{10}$. The following can be verified by inspection.

$$
(47.1)\qquad
\begin{aligned}
4 \cos a \varphi_5 &= \alpha_1(a) + \alpha_1'(a)\sqrt{5}, \\
4 \cos 2a\varphi_5 &= \alpha_2(a) + \alpha_2'(a)\sqrt{5}, \\
4 \cos (2a-1)\varphi_5 &= \alpha_3(a) + \alpha_3'(a)\sqrt{5}, \\
\sin (2a-1)\varphi_5 &= \alpha_4(a) \sin \varphi_5 + \alpha_4'(a) \sin 2\varphi_5;
\end{aligned}
$$

$$
\begin{aligned}
\alpha_1(a) &= ((1\ 1)_2\, 4\, (1\ 1)_2\, 4)_{10}, & \alpha_1'(a) &= (1_2\, \bar{1}_2\, 0\, \bar{1}_2\, 1_2\, 0)_{10}, \\
\alpha_2(a) &= (\bar{1}_4\, 4)_5, & \alpha_2'(a) &= (1\ 1_2\ 1\ 0)_5, \\
\alpha_3(a) &= (1_2\, 4\, 1_2)_5, & \alpha_3'(a) &= (1\ 1\ 0\ \bar{1}\ 1)_5, \\
\alpha_4(a) &= (1\ 0_3\ 1)_5, & \alpha_4'(a) &= (0\ 1\ 0\ 1\ 0)_5.
\end{aligned}
$$

$$
(47.2)\qquad
\begin{aligned}
4 \cos a\varphi_{10} &= \alpha_5(a) + \alpha_5'(a)\sqrt{5} + \alpha_5''(a) \cos \varphi_{10} + \alpha_5'''(a) \cos 3\varphi_{10}, \\
\cos (2a-1)\varphi_{10} &= \alpha_6(a) \cos \varphi_{10} + \alpha_6'(a) \cos 3\varphi_{10}, \\
4 \sin (2a-1)\varphi_{10} &= \alpha_7(a) + \alpha_7'(a)\sqrt{5};
\end{aligned}
$$

$$\alpha_5(a) = ((0\,1\,0\,\bar{1})_2\,0\,\bar{4}\,(0\,\bar{1}\,0\,1)_2\,0\,4)_{20}, \quad \alpha_5'(a) = (0\,(1\,0)_2\,(\bar{1}\,0)_2\,0_2\,(\bar{1}\,0)_2\,(1\,0)_2\,0)_{20},$$

$$\alpha_5''(a) = (4\,0_7\,\bar{4}\,0\,\bar{4}\,0_7\,4\,0)_{20}, \quad \alpha_5'''(a) = (0_2\,4\,0_3\,\bar{4}\,0_5\,\bar{4}\,0_3\,4\,0_3)_{20},$$

$$\alpha_6(a) = (1\,0_3\,\bar{1}_2\,0_3\,1)_{10}, \quad \alpha_6'(a) = (0\,1\,0\,\bar{1}\,0_2\,\bar{1}\,0\,1\,0)_{10},$$

$$\alpha_7(a) = (\bar{1}\,1\,4\,(1\,\bar{1})_2\,\bar{4}\,\bar{1}\,1)_{10}, \quad \alpha_7'(a) = (1_2\,0\,\bar{1}_2\,\bar{1}_2\,0\,1_2)_{10}.$$

It is important to observe that $K(\sqrt{5})$, $K(\cos\varphi_{10})$, $K(\cos 3\varphi_{10})$ are all different, and hence in a linear relation with rational coefficients between $\sqrt{5}$, $\cos\varphi_{10}$, $\cos 3\varphi_{10}$, coefficients of these, and the rational terms, may be separately equated. Similarly for $K(\sin\varphi_5)$, $K(\sin 2\varphi_5)$. For the reduction of non-linear functions of these irrationals the following will be found sufficient.

$$\sqrt{5}\cos\varphi_{10} = 2\cos 3\varphi_{10} + \cos\varphi_{10}, \quad \sqrt{5}\sin\varphi_5 = 2\sin 2\varphi_5 - \sin\varphi_5.$$

For the Pell case we find

$$2\cos a\varphi_8 = \beta_1(a) + \beta_1'(a)\sqrt{2} + \beta_1''(a)\cos\varphi_8 + \beta_1'''(a)\cos 3\varphi_8,$$

(48.1) $$\cos(2a-1)\varphi_8 = \beta_2(a)\cos\varphi_8 + \beta_2'(a)\cos 3\varphi_8,$$

$$\sin(2a-1)\varphi_8 = \beta_3(a)\sin\varphi_8 + \beta_3'(a)\sin 3\varphi_8;$$

$$\beta_1(a) = (0_7\,\bar{2}\,0_7\,2)_{16}, \quad \beta_1'(a) = (0\,1\,0_3\,(1\,0_3)_2\,1\,0_2)_{16},$$

$$\beta_1''(a) = (2\,0_5\,\bar{2}\,0\,2\,0_5\,\bar{2}\,0)_{16}, \quad \beta_1'''(a) = (0_2\,2\,0\,\bar{2}\,0_5\,2\,0\,2\,0_3)_{16},$$

$$\beta_2(a) = (1\,0_2\,\bar{1}_2\,0_2\,1)_8, \quad \beta_2'(a) = (0\,1\,1\,0_2\,\bar{1}\,\bar{1}\,0)_8,$$

$$\beta_3(a) = (1\,0_2\,\bar{1}\,1\,0_2\,\bar{1})_8, \quad \beta_3'(a) = (0\,1_2\,0_2\,\bar{1}_2\,0)_8.$$

$K(\sqrt{2})$, $K(\cos\varphi_8)$, $K(\cos 3\varphi_8)$ are distinct, also $K(\sin\varphi_8)$, $K(\sin 3\varphi_8)$. For the Lucas case

$$4\cos a\varphi_{12} = \gamma_1(a) + \gamma_1'(a)\sqrt{2} + \gamma_1''(a)\sqrt{3} + \gamma_1'''(a)\sqrt{6},$$

(49.1) $$4\cos(2a-1)\varphi_{12} = \gamma_2(a)\sqrt{2} + \gamma_2'(a)\sqrt{6},$$

$$4\sin(2a-1)\varphi_{12} = \gamma_3(a)\sqrt{2} + \gamma_3'(a)\sqrt{6};$$

$$\gamma_1(a) = (0_3\,2\,0_3\,\bar{2}\,0_3\,\bar{4}\,0_3\,\bar{2}\,0_3\,2\,0_3\,4)_{24},$$

$$\gamma_1'(a) = (1\,0\,2\,0\,\bar{1}\,0\,1\,0\,2\,(0\,1)_2\,0\,\bar{2}\,0\,1\,0\,1\,0\,2\,0\,1\,0)_{24},$$

$$\gamma_1''(a) = (0\,2\,0_7\,\bar{2}\,0_3\,\bar{2}\,0_7\,2\,0_2)_{24},$$

$$\gamma_1'''(a) = (1\,0_3\,1\,0\,\bar{1}\,0_3\,\bar{1}\,0\,\bar{1}\,0_3\,\bar{1}\,0\,1\,0_3\,1\,0)_{24}.$$

$$\gamma_2(a) = (1\,2\,\bar{1}\,1\,\bar{2}\,\bar{1}_2\,2\,1\,\bar{1}\,2\,1)_{12}, \quad \gamma_2'(a) = (1\,0\,1\,\bar{1}\,0\,\bar{1}_2\,0\,1\,1\,0\,1)_{12},$$

$$\gamma_3(a) = (\bar{1}\,2\,\bar{1}_2\,2\,\bar{1}\,1\,\bar{2}\,\bar{1}_2\,\bar{2}\,1)_{12}, \quad \gamma_3'(a) = (1\,0\,1_2\,0\,1\,1\,0\,1_2\,0\,1)_{12}.$$

(51.1) $\cos a\varphi_1 = (-1)^a, \quad \cos(2a-1)\varphi_2 = 0, \quad \sin(2a-1)\varphi_2 = (-1)^{a-1}.$

For the Fermat and miscellaneous cases

$$2\cos a\varphi_3 = \delta_1(a) = (1\ \bar{1}\ \bar{2}\ 1\ 1\ 2)_6,$$
$$2\cos 2a\varphi_3 = \delta_2(a) = (\bar{1}_2\ 0)_3,$$
$$2\cos(2a-1)\varphi_3 = \delta_3(a) = (1\ \bar{2}\ 1)_3,$$
$$2\sin(2a-1)\varphi_3 = \delta_4(a)\sqrt{3}, \qquad \delta_4(a) = (1\ 0\ \bar{1})_3.$$
$$2\cos a\varphi_4 = \varepsilon_1(a) + \varepsilon_1'(a)\sqrt{2}, \quad \varepsilon_1(a) = (0_3\ \bar{2}\ 0_2\ 1\ 2)_8, \quad \varepsilon_1'(a) = (1\ 0\ \bar{1}\ 0\ \bar{1}\ 0)_8,$$
$$2\cos(2a-1)\varphi_4 = \varepsilon_2(a)\sqrt{2}, \qquad \varepsilon_2(a) = (1\ \bar{1}_2\ 1)_4,$$
$$2\sin(2a-1)\varphi_4 = \varepsilon_3(a)\sqrt{2}, \qquad \varepsilon_3(a) = (1_2\ \bar{1}_2)_4.$$
$$2\cos a\varphi_6 = \zeta_1(a) + \zeta_1'(a)\sqrt{3},$$
$$\zeta_1(a) = (0\ 1\ 0\ 1\ 0\ \bar{2}\ 0\ 1\ 0\ 1\ 0\ 2)_{12}, \qquad \zeta_1'(a) = (1\ 0_3\ \bar{1}\ 0\ \bar{1}\ 0_3\ 1\ 0)_{12},$$
$$2\cos(2a-1)\varphi_6 = \zeta_2(a)\sqrt{3}, \qquad \zeta_2(a) = (1\ 0\ \bar{1}_2\ 0\ 1)_6,$$
$$2\sin(2a-1)\varphi_6 = \zeta_3(a)\sqrt{3}, \qquad \zeta_3(a) = (1\ 2\ 1\ \bar{1}\ \bar{2}\ \bar{1})_6.$$

It is possible to state $\alpha, \beta, \cdots, \zeta$ as explicit functions of their arguments by means of Legendre-Jacobi symbols; in all but a few cases the results are too complicated to be of interest.

16. **Fibonacci relations.** These are written out by applying (47), (47.1) (47.2) to the identities in § 5 for $x = \varphi_5$ or φ_{16}. For these values (4.1) gives

$$\text{(XIV)} \qquad \begin{aligned} \sum v_r O_{2r}(2n) &= 2P(n) + \sum \alpha_2(a) P(n-2a^2), \\ \sum u_r O_{2r}(2n) &= \sum \alpha_2'(a) P(n-2a^2), \end{aligned}$$
$$\sum 5^r [v_{2r} O_{4r}(2n) + 5u_{2r+1} O_{4r+2}(2n)] = 2P(n) + \sum \alpha_1(a) P(n-2a^2),$$
$$\sum 5^r [u_{2r} O_{4r}(2n) + v_{2r+1} O_{4r+2}(2n)] = \sum \alpha_1'(a) P(n-2a^2).$$

Similarly from (5.1),

$$\text{(XV)} \qquad \begin{aligned} \sum v_{2r+1} O_{2r+1}(2n+1) &= \sum \alpha_3(a) P(n-4t_a), \\ \sum u_{2r+1} O_{2r+1}(2n+1) &= \sum \alpha_3'(a) P(n-4t_a), \end{aligned}$$
$$\sum 5^r [u_{2r+1} O_{4r+1}(2n+1) + v_{2r+2} O_{4r+3}(2n+1)] = \sum \alpha_8(a) P(n-4t_a),$$
$$\sum 5^r [v_{2r+1} O_{4r+1}(2n+1) + u_{2r+1} O_{4r+3}(2n+1)] = \sum \alpha_8'(a) P(n-4t_a),$$
$$\alpha_8(a) = (1\ 0_3\ \bar{1}_2\ 0_3\ 1)_{10}, \qquad \alpha_8'(a) = (1\ 2\ 0\ \bar{2}\ \bar{1}_2\ \bar{2}\ 0\ 2\ 1)_{10}.$$

From (6.1),

$$\sum(-1)^r u_r A_r(n) = \sum \alpha_9(a) P(n - t_a),$$

$$\sum(-1)^r v_r A_r(n) = \sum \alpha_9'(a) P(n - t_a),$$

(XVI) $$\alpha_9(a) = (0\ 1\ 2\ \bar{1}\ 0)_5, \qquad \alpha_9'(a) = (2\ 3\ 2\ 3\ 2)_5,$$

$$\sum u_r A_r(n) = \sum \alpha_6(a) P(n - t_a),$$

$$\sum v_r A_r(n) = \sum \alpha_{10}(a) P(n - t_a),$$

$$\alpha_{10}(a) = (2\ 1\ 0\ 1\ 2_2\ 1\ 0\ \bar{1}\ 2)_{10}.$$

From (6.2),

$$5 A_0(n) + \sum(2 v_a + v_{a+1}) A_a(n) = 5 P(n) + 5 \sum \alpha_{11}(a) P(n - t_{a+1}),$$

(XVII) $$A_0(n) + \sum(2 u_a + u_{a+1}) A_a(n) = P(n) + \sum \alpha_{11}'(a) P(n - t_{a+1}),$$

$$\alpha_{11}(a) = (0_3\ 1_2\ 0_3\ 1_2)_{10}, \qquad \alpha_{11}'(a) = (2\ 0_2\ 1_2\ \bar{2}\ 0_2\ 1_2)_{10};$$

write for brevity $B_0(n) = A_0(n)$, $B_a(n) = A_{a-1}(n) + 2 A_a(n)$;

$$4 B_0(n) + \sum 5^a [v_{2a} B_{4a}(n) + u_{2a-1} B_{4a-2}(n)]$$
$$= 4 P(n) + \sum \alpha_{12}(a) P(n - t_{a+1}),$$

$$\sum 5^{a-1} [5 u_{2a} B_{4a}(n) + v_{2a-1} B_{4a-2}(n)] = \sum \alpha_{12}'(a) P(n - t_{a+1}),$$

$$\sum 5^{a-1} [(v_{2a-1} + 5 u_{2a-1}) B_{4a-1} + (v_{2a-2} + u_{2a-2}) B_{4a-3}]$$
$$= 2 P(n) + \sum \alpha_{12}''(a) P(n - t_{a+1}),$$

$$\sum 5^{a-1} [v_{2a-1} B_{4a-1} + u_{2a-2} B_{4a-3}] = \sum \alpha_{12}'''(a) P(n - t_{a+1}),$$

$$\alpha_{12}(a) = (1_2\ \bar{1}_2\ 1_2\ 4_2\ \bar{1}_2\ 1_2\ 1_2\ 1_2\ 4_2)_{20}, \qquad \alpha_{12}'(a) = (1_4\ 1_4\ 0_2\ 1_4\ 1_4\ 0_2)_{20},$$

$$\alpha_{12}''(a) = (2\ 0_6\ \bar{2}_4\ 0_6\ 2_3)_{20}, \qquad \alpha_{12}'''(a) = (0\ 1_2\ 0_2\ 1_2\ 0_4\ 1_2\ 0_2\ 1_2\ 0_3)_{20}.$$

From (4.11),

$$\sum 5^r [v_{2r} O_{4r}(2n) + 5 u_{2r+1} O_{4r+2}(2n)]$$
$$= 2 P(n) + \sum(-1)^a \alpha_2(a) P(n - 2a^2),$$

(XVIII) $$\sum 5^r [u_{2r} O_{4r}(2n) + v_{2r+1} O_{4r+2}(2n)]$$
$$= -\sum(-1)^a \alpha_2'(a) P(n - 2a^2);$$

$$\sum v_{2r} O_{2r}(2n) = 2 P(n) + \sum(-1)^a \alpha_1(a) P(n - 2a^2),$$

$$\sum u_{2r} O_{2r}(2n) = -\sum(-1)^a \alpha_1'(a) P(n - 2a^2).$$

From (5.11),

$$\sum 5^r \left[u_{2r} O_{4r+1}(2n+1) + v_{2r+1} O_{4r+3}(2n+1)\right] = \sum (-1)^a \alpha_4'(a)\, P(n-4t_a),$$

$$\text{(XIX)} \quad \sum 5^r \left[v_{2r} O_{4r+1}(2n+1) + 5u_{2r+1} O_{4r+3}(2n+1)\right] = \sum (-1)^a \alpha_{13}(a)\, P(n-4t_a),$$

$$\alpha_{13}(a) = (\bar{1}_2\ 0\ 1_2)_5;$$

$$\sum v_{2r} O_{2r+1}(2n+1) = -\sum (-1)^a \alpha_7(a)\, P(n-4t_a),$$

$$\sum u_{2r} O_{2r+1}(2n+1) = \sum (-1)^a \alpha_7'(a)\, P(n-4t_a).$$

From (6.11),

$$\sum u_r A_r(n) = 2\sum (-1)^a \alpha_4'(a)\, P(n-t_a),$$

$$\sum v_r A_r(n) = 2\sum (-1)^a \alpha_{14}(a)\, P(n-t_a), \qquad \alpha_{14}(a) = (1\ 2\ 0\ \bar{2}\ \bar{1})_5;$$

$$\text{(XX)} \quad \sum (-1)^r v_r A_r(n) = \sum (-1)^a \alpha_{15}'(a)\, P(n-t_a),$$

$$\sum (-1)^r u_r A_r(n) = \sum (-1)^a \alpha_{15}(a)\, P(n-t_a).$$

$$\alpha_{15}(a) = (1\ 0\ \bar{2}\ 0\ 1\ \bar{1}\ 0\ 2\ 0\ 1)_{10}, \qquad \alpha_{15}'(a) = (3\ 2\ \bar{2}\ 2\ 3\ \bar{3}\ 2\ \bar{2}\ \bar{2}\ 3)_{10}.$$

(6.21) gives nothing new.

17. **Pell, Lucas, Fermat relations.** Similarly to § 16 the Pell relations are written out from (48), (48.1); the Lucas from (49), (49.1) and the Fermat from (50), (51), (51.1). To save space we omit the results.

18. **Miscellaneous and aperiodic relations.** On account of their simplicity we include a few relations found by applying certain of the formulas in (50), (51.1) to the fundamental identities. In (4.1), (5.1) put $x = \varphi_3$, and in (6.1) $x = \varphi_6$:

$$O(2n) = P(n) + \sum \delta_2(a)\, P(n-2a^2),$$

$$\text{(XXI)} \quad O(2n+1) = \sum \delta_3(a)\, P(n-4t_a),$$

$$A(n) = \sum \zeta_2(a)\, P(n-t_a),$$

which express the O, A functions in terms of P's.

The next are found in a wholly different way from any of the preceding. In the fundamental identities (4)–(6.21) reduce all powers of cosines or sines on the left to functions of multiple angles and in the results equate coefficients of cosines or sines having equal arguments; or express all cosines or sines of multiple angles on the right of the original identities in terms of powers of $\cos x$, $\sin x$, and in these results equate coefficients of like powers of $\cos x$ or $\sin x$.

Equating first coefficients of $\cos 2sx$ then of $\cos^{2s} x$, $s \geqq 0$, in (4.1) we find

$$\text{(XXII)}\quad \begin{aligned} P(n) &= \sum \binom{2s+2r}{r} O_{2s+2r}(2n+4s^2), \\ O_{2s}(2n) &= P(n-2s^2) + 2\sum (-1)^a \frac{s+a}{2s+a} \binom{2s+a}{a} P(n-2(s+a)^2). \end{aligned}$$

The first expresses $P(n)$ linearly in terms of O-partitions of ranks $2s$, $2(s+1)$, $2(s+2)$, $\cdots$ of the fixed integer $2n+4s^2$; the second gives the O-partitions of $2n$ of fixed rank $2s$ in terms of all the partitions of $n-2s^2$, $n-2(s-1)^2$, $n-2(s-2)^2$, $\cdots$. From their derivations these are valid for $s \geqq 0$. Similarly from (5.1),

$$\text{(XXIII)}\quad \begin{aligned} P(n) &= \sum \binom{2s+2r+1}{r} O_{2s+2r+1}(2n+(2s+1)^2), \\ O_{2s+1}(2n+1) &= \sum (-1)^{a-1} \frac{2s+2a-1}{2s+a} \binom{2s+a}{a-1} P(n-4t_{s+a}). \end{aligned}$$

We omit the corresponding pair for $A_a(n)$, $P(n)$ found in the same way from (6.2). By combining Euler's expansion of $\prod (1-x^n)$ with the simplest identities between the theta functions it is easy to derive recurrences for $A(n)$, $O(n)$ in which the arguments decrease by multiples of the successive triangular numbers.

* University of Washington.
December, 1921.

FUNCTIONALS OF SUMMABLE FUNCTIONS.*

By William L. Hart.

Introduction. Let H represent the class of all real valued functions $u(x)$, defined on an interval $a \leq x \leq b$ which, together with their squares $u^2(x)$, are summable in the Lebesgue sense on (a, b). In the present paper we shall obtain certain results concerning functionals whose arguments are functions $u(x)$ of H. Functionals of this type have been considered previously† but the main results of the present paper are new and the point of view throughout is different from that of previous authors. In many instances below, the Riesz-Fischer theorem‡ concerning the Fourier constants of a summable function is appealed to in order to reduce questions concerning functionals defined in H to related questions concerning functions of infinitely many variables defined in Hilbert space.

In Part I of the paper we shall consider functionals $F[u]$ which are continuous and possess differentials according to customary definitions. For the general linear functional, a representation by means of an infinite series is obtained. A mean value theorem is established for $F[u]$ and an infinite system of functional equations is solved in which the functionals involved are of type $F[u]$.

In Part II of the paper we shall discuss functionals $G[u, s]$ which, for every u of H, are summable functions of s defined on an interval $c \leq s \leq d$. In reference to $G[u, s]$, appropriate definitions are given for the terms *continuity* and *differential*. A representation for the general linear functional is obtained which involves an infinite series converging in the mean. The functions of infinitely many variables related to functionals $G[u, s]$ are shown to have partial *pseudo-derivatives* with respect to their arguments in the sense previously defined§ by the author. A mean value theorem for functionals $G[u, s]$ is obtained and a type of implicit functional equations is solved. A differential-functional equation is also considered where the derivative entering is a pseudo-derivative.

* Presented to the American Mathematical Society at Chicago, Dec. 30, 1920.

† Lévy, Bulletin de la Société Mathématique de France, vol. XLVIII (1920), p. 13.

‡ Cf. Plancherel, Rendiconti del Circolo Matematico di Palermo, vol. 30 (1910), p. 296.

§ Bulletin of the American Mathematical Society, vol. XXVII (1921), p. 202. Referred to in the future as Paper I.

All integrals below are taken in the sense of Lebesgue. A function $f(x)$ will be termed *integrable* if both f and f^2 are summable in the Lebesgue sense. All functions and variables will be supposed real valued. The phrase *almost everywhere* will mean *with the exception at most of a set of points of measure zero.* In the discussion below two functions $h_1(x)$ and $h_2(x)$, defined on a set E, well be called the same if they are equal almost everywhere on E. In all equations, as for example $h_1(x) = h_2(x)$, it will be understood that the equality may fail to hold on some sub-set of E of measure zero.

If $u(x)$ is a function of H we shall denote the positive square root of $\int_{(ab)} u^2(x)\,dx$ by Mu and shall call it the modulus of $u(x)$. We shall have occasion many times to refer to a function $u(x)$ of H as the *point* $u(x)$, and we may think of the modulus of $u(x)$ as representing the distance of $u(x)$ from the origin, $u = 0$, in H.

Let I represent a sequence of integrable functions $[p_i(x); i = 1, 2, \cdots]$, which are unitary and orthogonal on the interval (a, b) and which, moreover, form a complete set for the class H. That is, there does not exist in H any function $u(x) \neq 0$ for which $z_i = 0$ $(i = 1, 2, \cdots)$, where

$$(1) \qquad z_i = \int_{(ab)} u(x)\,p_i(x)\,dx.$$

We shall call the numbers z_i the *Fourier coefficients* of $u(x)$ and shall term $\zeta = (z_1, z_2, \cdots)$ the *Fourier vector* corresponding to $u(x)$. It is well known that $(Mu)^2 = \sum_{i=1}^{\infty} z_i^2$ and, accordingly, we shall call the positive square root of this infinite sum the modulus of ζ and shall represent it by $M\zeta$. Since $M\zeta$ exists, the vector ζ may be considered as representing a point in real Hilbert space of infinitely many dimensions. Moreover, as a consequence of the Riesz-Fischer theorem, it is known that a one to one correspondence exists between points $u(x)$ in H and points ζ in Hilbert space. If ζ is the point corresponding to a function $u(x)$ we shall denote this fact by the notation $\zeta \rightleftarrows u(x)$.

It will be useful later to recall that, if $u_n = z_1 p_1 + \cdots + z_n p_n$, then

$$(2) \qquad \lim_{n=\infty} M^2(u_n - u) = \lim_{n=\infty} \sum_{i=n+1}^{\infty} z_i^2 = 0. \qquad (\zeta \rightleftarrows u(x))$$

Moreover, if $(u_n(x); n = 1, 2, \cdots)$ is a sequence of H satisfying

$$(3) \qquad \lim_{n=\infty} M(u_n - u) = 0,$$

it is easily verified that

$$(4) \qquad \lim_{n=\infty} M(\zeta_n - \zeta) = 0. \qquad (\zeta_n \rightleftarrows u_n;\ \zeta \rightleftarrows u)$$

When a sequence (u_n) satisfies (3) it is customary to say that u is the *limit in the mean** of the sequence (u_n). We shall use the notation

(5) $$\underset{n=\infty}{\text{l. m.}}\, u_n = u,$$

to denote such convergence. Similarly, we shall abbreviate the convergence in (4) by the notation $\underset{n=\infty}{\text{l. m.}}\, \zeta_n = \zeta$. If the partial sums $S_n(x)$ of a series $\sum_{i=1}^{\infty} u_i(x)$ of functions of H satisfy the equation $\underset{n=\infty}{\text{l. m.}}\, S_n(x) = S(x)$, we shall say that the series converges in the mean to $S(x)$ and we shall indicate this fact by the notation,

(6) $$\sum_{i=1}^{\infty}{}^{(m)}\, u_i(x) = S(x).$$

Part I.

Functionals without a Parameter.

1. Definitions and theorems on linear functionals. Let $F[u]$ be a functional operation which, for every u in H, yields a real number. Then we shall say that $F[u]$ is continuous† at u if, whenever equation (5) is satisfied, it follows that $\lim_{n=\infty} F[u_n] = F[u]$. It is seen immediately that, if $u_1(x) = u_2(x)$ almost everywhere, then $F[u_1] = F[u_2]$. In the future all functionals considered will be supposed continuous.

A functional $L[u]$ is linear if, for every pair of points u_1 and u_2 in H, $L[u_1 + u_2] = L[u_1] + L[u_2]$. It is easily verified that, for every real constant c and for every u of H, $L[cu] = cL[u]$.

Definition 1.‡ *Suppose that, corresponding to a point u_0 in H, there exists a linear functional $L[u_0, v]$, defined for all functions $v(x)$ in H and satisfying the equation*

(7) $$F[u] = F[u_0] + L[u_0, v] + (Mv)\,\epsilon(u_0, v),$$

where $u = u_0 + v$ and where $\epsilon(u_0, v)$, approaches zero with Mv. Then, we shall call $L[u_0, v]$ the differential of F at the point u_0.

It is obvious that the differential is unique, if it exists, because, as a consequence of (7),

$$L[u_0, v] = \lim_{d=0} \frac{F[u_0 + dv] - F[u_0]}{d}.$$

* Cf. Plancherel, loc. cit., p. 292.

† Cf. Lévy, loc. cit., p. 21.

‡ Cf. Lévy, loc. cit., p. 21.

Let $f(\zeta)$ be a function whose value, for every ζ in Hilbert space, is defined by the equation $f(\zeta) = F[u]$ where $u = \zeta$. We shall say that $f(\zeta)$ *corresponds to* $F[u]$ *and shall* denote this fact by the notation $f(\zeta) \sim F[u]$. In view of equations (3) and (4) it is seen that $f(\zeta)$ is continuous in the weak* sense at every point ζ in Hilbert space.

THEOREM 1. *If* $L[u]$ *is a linear functional, then*

$$L[u] = \sum_{i=1}^{\infty} z_i L[p_i]. \qquad (\zeta = u) \tag{8}$$

As a consequence of the linearity and the continuity of $L[u]$, it follows that

$$L[u] = \lim_{n=\infty} L[z_1 p_1 + \cdots + z_n p_n] = \lim_{n=\infty} \sum_{i=1}^{n} z_i L[p_i],$$

which establishes the theorem.

Since (8) must converge for all points ζ in Hilbert space, we may state as an immediate corollary that the series

$$\sum_{i=1}^{\infty} (L[p_i])^2 \tag{9}$$

converges.† Moreover, it is seen that the function $l(\zeta)$, corresponding to $L[u]$, is given by the right member of (8) and hence is a linear form in infinitely many variables.

Corollary 1.‡ *If* $L[u]$ *is linear, there exists a constant* $K > 0$ *such that, for all* u *of* H, $|L[u]| \leq KMu$.

If we apply the Schwarz inequality for sums in (8) we obtain

$$|L[u]| \leq (Mu) \Big[\sum_{i=1}^{\infty} (L[p_i])^2 \Big]^{\frac{1}{2}},$$

from which the corollary follows with an obvious definition for K.

If we insert expressions (1) in (8), we immediately obtain the equation

$$L[u] = \lim_{n=\infty} \int_{(ab)} K_n(x)\, u(x)\, dx \qquad (K_n(x) = p_1(x) L[p_1] + \cdots + p_n(x) L[p_n])$$

It should be noted that the functions (K_n) do not depend on u.

* Cf. Hart, Transactions of the American Mathematical Society, vol. XXIII (1922), p. 32. This paper will be referred to in the future as Paper II.

† Cf. Riesz, *Équationes Linéaires*, p. 47.

‡ Cf. Lévy, loc. cit., p. 20.

THEOREM 2. *Suppose that $F[u]$ possesses a differential $L[u, v]$ at the point u in H. Then, if $f(\zeta) \equiv F[u]$, it follows that, at the point ζ in Hilbert space corresponding to u, there exist partial derivatives*

$$\text{(10)} \qquad \frac{\partial f(\zeta)}{\partial z_i} = L[u, p_i]. \qquad (i = 1, 2, \cdots)$$

Moreover, the following series converges:

$$\text{(11)} \qquad \sum_{i=1}^{\infty}\left(\frac{\partial f(\zeta)}{\partial z_i}\right)^2.$$

In order to establish (10) for the case $i = 1$, for example, we start from the equation

$$\begin{aligned} f(z_1 + d, z_2, z_3, \cdots) - f(\zeta) &= F[u + d p_1] - F[u], \\ &= L[u, d p_1] + |d|(M p_1) e(u, d p_1), \\ &= d L[u, p_1] + |d|(M p_1) e(u, d p_1). \end{aligned}$$

Hence, since $\lim\limits_{d=0} M(d p_1) = 0$, it follows that (10) is true for $i = 1$. A similar proof establishes the result for all values of i. The convergence of (11) is a direct consequence of (9).

The conclusions of Theorem (1), applied to the differential $L[u; v]$, permit us to state a corresponding result for $F[u]$.

THEOREM 3. *Under the hypotheses of Theorem 2, it follows that*

$$F[u'] = F[u] + \sum_{i=1}^{\infty}(z_i' - z_i) L[u, p_i] + (M v) e(u, v),$$

where $v = u' - u$ and where $u' \equiv \zeta' = (z_1', z_2', \cdots)$.

2. A mean value theorem and the solution of functional equations. With the aid of the results of the previous section we may obtain several theorems in regard to functionals by the use of theorems previously proved by the author* for functions defined in Hilbert space.

Let $T[u, v]$ be a functional defined for all (u, v) in H. We shall say that T is continuous simultaneously in its arguments if the equations

$$\underset{n=\infty}{\text{l. m.}}\, u_n = u, \qquad \underset{n=\infty}{\text{l. m.}}\, v_n = v,$$

where (u_n, v_n, u, v) are in H, imply that $\lim\limits_{n=\infty} T[u_n, v_n] = T[u, v]$.

* Cf. Paper II.

THEOREM 4. *Let $F[u]$ possess a differential $L[u, v]$ for all u in H and suppose that L is continuous simultaneously in its arguments for all (u, v) in H. Assume, moreover, that a constant $K > 0$ exists such that, for all (u, v),*

$$L[u, v] \leqq KMv. \tag{13}$$

Then it follows that, if (u', u) are in H,

$$F[u'] - F[u] = \int_0^1 L[u + tv, v]\, dt \qquad (v = u' - u) \tag{14}$$

$$= \sum_{i=1}^{\infty} (z_i' - z_i) \int_0^1 L[u + tv, p_i]\, dt. \qquad (u = \zeta;\ u' = \zeta') \tag{15}$$

Since L is linear in v, we know that, for every u, a constant $K(u)$ exists satisfying (13). The force of the assumption (13) lies in the supposition that $K(u) < K$ for all u. By virtue of (13), since L is a linear form in the quantities $(z_i' - z_i)$, a well known theorem* permits us to state that, for every u,

$$\sum_{i=1}^{\infty} (L[u, p_i])^2 \leqq K^2. \tag{16}$$

Suppose, now, that two points (u', u) in H are assigned and let us establish (14). It is seen that, for a given pair (u, v), the function of r defined by $S(r) = F[u + rv]$ is continuous and has a continuous derivative $L[u + rv, v]$ for $0 \leqq r \leqq 1$. Consequently, the mean value theorem in integral form, applied to the difference $S(1) - S(0)$, establishes (14).

In considering (15) we first note that, because of (16), the series

$$\sum_{i=1}^{\infty} (z_i' - z_i)\, L[u + tv, p_i] \tag{17}$$

converges uniformly for $0 \leqq t \leqq 1$. Since the integrand in (14) is equal to (17), it follows that the right member of (15) exists and is equal to that of (14).

The theorem could also have been proved by an application of the mean value theorem for functions in Hilbert space.†

Let us consider the infinite system of equations

$$F_i[u, t] = 0, \qquad (i = 1, 2, \cdots) \tag{18}$$

where, for every t on an interval $c < t \leqq d$, $F_i[u, t]$ is a functional of the type F treated previously. Suppose that, for every t, $F_i[u, t]$ possesses

* Cf. Hellinger and Toeplitz, Mathematische Annalen, vol. LXIX (1910), p. 295.

† Paper II, Theorem II.

a differential $L_i[u, v; t]$ at every point u in H, and assume that both F_i and L_i are continuous in their arguments simultaneously. That is, suppose that the equations

$$\underset{n=\infty}{\mathrm{l.m.}}\, u_n = u, \qquad \underset{n=\infty}{\mathrm{l.m.}}\, v_n = v, \qquad \lim_{n=\infty} (t_n - t) = 0,$$

where (u_n, v_n, u, v) are in H and where (t_n, t) are on (c, d), imply that

$$\lim_{n=\infty} F_i[u_n, t_n] = F_i[u, t], \qquad \lim_{n=\infty} L_i[u_n, v_n; t_n] = L_i[u, v; t].$$

Under these assumptions it is seen that, if $f_i(\zeta, t) \equiv F_i[u, t]$ and $l_i(\zeta, v; t) = L_i[u, v; t]$, then

$$\frac{\partial f_i(\zeta, t)}{\partial z_j} = L_i[u, p_j; t] = l_i[\zeta, p_j; t]. \qquad (i, j = 1, 2, \cdots; \zeta \equiv u) \tag{19}$$

Therefore, the derivatives (19) are continuous in their arguments simultaneously, in the weak sense with respect to ζ and in the ordinary sense with respect to t.

The previous theorems of this paper together with the theorem* on implicit functions for Hilbert space applied to the system

$$f_i(\zeta, t) = 0, \qquad (i = 1, 2, \cdots)$$

permit us to state the following theorem without further proof.

THEOREM 5. *Assume that the series*

$$\sum_{i=1}^{\infty} F_i^2[u, t], \quad \sum_{\substack{i,j=1 \\ i \neq j}}^{\infty} L_i^2[u, p_j; t], \quad \sum_{i=1}^{\infty} (L_i[u, p_i; t] - 1)^2, \tag{20}$$

converge uniformly for all u in H and for $c \leq t \leq d$. Suppose, moreover, that the infinite determinant

$$\Delta = |L_i[u_0, p_j; t]|_{i,j=1,2,\cdots} \tag{21}$$

is not zero. Then, if $(u = u_0, t = t_0)$ is a solution of (18) we can select a number $r > 0$ so small that, for $|t - t_0| \leq r$, there exists uniquely a function $u(x, t)$ satisfying (18). For every t on $|t - t_0| \leq r$ the function $u(x, t)$ is in H. The coordinates $z_i(t)$ of the point $\zeta(t) \equiv u(x, t)$ are continuous functions of t and

$$\lim_{t=t_0} M\zeta(t) = \lim_{t=t_0} Mu(x, t) = Mu_0. \tag{22}$$

* Cf. Paper II, Theorem V.

PART II.

Functionals Containing a Parameter.

1. Definitions and theorems on linear functionals. Let W represent the class of all functions $w(x)$ which are integrable on an interval $E = (c \leqq s \leqq d)$. Suppose that $G[u, s]$ is a functional operation which, for every u of H, yields a function of s belonging to the class W.

Definition 1.* *The functional $G[u, s]$ is continuous at u if, whenever* $\underset{n=\infty}{\text{l. m.}}\, u_n = u$, *it follows that*

$$\underset{n=\infty}{\text{l. m.}}\, G[u_n, s] = G[u, s]. \tag{23}$$

The functional is linear if $G[u_1, s] + G[u_2, s] = G[u_1 + u_2, s]$ for all (u_1, u_2) in H.

It should be recalled that the statement (23) implies that

$$\lim_{n=\infty} \int_E (G[u_n, s] - G[u, s])^2\, ds = 0.$$

All functionals considered below will be supposed continuous. It is seen that, if $u_1(x) = u_2(x)$ almost everywhere, then $G[u_1, s] = G[u_2, s]$. Moreover, if $L[u, s]$ is a linear functional, it is easily established that $L[cu, s] = cL[u, s]$ for every constant c and for u in H. It is known† that, for every linear functional $L[u, s]$, there exists a constant $K > 0$ such that

$$ML[u, s] = \sqrt{\int_E L^2[u, s]\, ds} \leqq KMu.$$

THEOREM 1. *If $L[u, s]$ is a linear functional, then, for every u in H,*

$$L[u, s] = \sum_{i=1}^{\infty} (m)\, z_i\, L[p_i, s], \qquad (\zeta = u) \tag{24}$$

$$= \underset{n=\infty}{\text{l. m.}} \int_{(ab)} V_n(x, s)\, u(x)\, dx, \tag{25}$$

where $V_n(x, s) = p_1(x)\, L[p_1, s] + \cdots + p_n(x)\, L[p, s]$.

The proof of (24) would make use of the continuity of L and would be similar in method to the proof of Theorem (1), Part I, with "limit in the mean" replacing "limit" as met in Part I. Equation (25) is obtained from (24) by inserting the expressions (1) for the z_i. It should be noted that $V_n(x, s)$ is independent of $u(x)$. Moreover, from a known property of convergence in

* Cf. Lévy, loc. cit., p. 20.

† Cf. Lévy, loc. cit., p. 20.

the mean,* it follows that a sub-sequence (V_{h_n}; $n = 1, 2, \cdots$) can be selected from the sequence (V_n) so that

$$\lim_{n=\infty} \int_{(ab)} V_{h_n}(x,s)\,u(x)\,dx = \lim_{n=\infty} \sum_{i=1}^{h_n} z_i L[p_i, s] = L[u, s],$$

almost everywhere on E.

Let $L[u_0; v, s]$, where u_0 is a point in H, be a linear functional defined for v in H and for s on E.

Definition 2.† *Suppose that the functional G satisfies the equation*

$$G[u, s] - G[u_0, s] = L[u_0; v, s] + (Mv)e(u_0, v, s), \qquad (v = u - u_0) \tag{26}$$

where $Me(u_0, v, s)$ approaches zero with Mv. Then, $L[u_0; v, s]$ is defined as the differential of $G[u, s]$ at the point u_0.

The uniqueness of the differential L, if it exists, is easily established by use of the notion of a pseudo-derivative as previously defined by the author.‡ Let $R(r, s)$ be integrable on E for every r on an interval T. Then, the pseudo-derivative of R with respect to r at the point $r = r_0$ is the integrable function $R_r(r_0, s)$ which satisfies the equation

$$\underset{h=0}{\mathrm{l.\,m.}} \frac{\Delta R}{h} = R_r(r_0, s), \qquad [\Delta R = R(r_0 + h, s) - R(r_0, s)] \tag{27}$$

provided that the limit exists. It should be recalled that (27) implies that

$$\lim_{h=0} \int_E \left(R_r(r_0, s) - \frac{\Delta R}{h} \right)^2 ds = 0.$$

Moreover, when the pseudo-derivative at a point r_0 exists, it is unique up to its values at a set of points s of measure zero.

THEOREM 2. *Suppose that G possesses a differential L at all points in H. Let $R(r, s) = G[u + rv, s]$, where $0 \leq r < 1$. Then, for every value of r there exists a pseudo-derivative $R_r(r, s) = L[u + rv; v, s]$.*

In order to establish the theorem let us first note that, if $\lim_{n=\infty} r_n = r$, then $\underset{n=\infty}{\mathrm{l.\,m.}}\, u + r_n v = u + rv$. Consider the expression

$$\frac{\Delta R}{h} = \frac{L[u + rv; rh, s]}{h} + \frac{M(hv)}{h} e(u + rv, hv, s), \quad [\Delta R = R(r + h, s) - R(r, s)]$$

$$= L[u + rv; v, s] + (Mv)e(u + rv, hv, s).$$

* Cf. Plancherel, loc. cit., p. 294.

† Cf. Lévy, loc. cit., p. 21.

‡ Paper I, Definition (3).

Consequently, as a result of the properties of $e(u, v, s)$, it follows that

$$\lim_{h=0} \int_E \left(\frac{\Delta R}{h} - L[u + rv; v, s] \right)^2 ds = 0,$$

which establishes the theorem.

When we place $r = 0$ in Theorem 2 we obtain $R_r(0, s) = L[u; v, s]$. Hence, since the pseudo-derivative is unique, it follows that the differential $L[u; v, s]$ is unique, if it exists.

2. Expressions for $G[u, s]$ in terms of its differential. The mean value theorem* for a function possessing a pseudo-derivative, applied to $R(r, s)$ enables us to obtain a mean value theorem for G without further proof.

THEOREM 3. *Let $G[u, s]$ satisfy all conditions of Theorem 2 and suppose that*

$$\lim_{r_1 = r} \int_E (L[u + r_1 v; v, s] - L[u + rv; v, s])^2 ds = 0,$$

for all values of r on $(0, 1)$. Assume that, for every pair (u, v), $L[u + rv; v, s]$ is integrable with respect to (r, s) in the rectangle $(0 \leq r \leq 1; c < s < d)$. Then it follows that, for every (u_1, u) in H,

$$G[u_1, s] - G[u, s] = \int_{(0,1)} L[u + rv; v, s] dr. \qquad (v = u_1 - u) \tag{28}$$

Since $L[u; v, s]$ is linear in v, there exists, for every u, a constant $K(u)$ statistying the equation

$$M(L[u; v, s]) < K(u) Mv.$$

Corollary 1. Assume that the hypotheses of Theorem 3 are satisfied and that a constant $K > 0$ exists such that $K(u) \leq K$ for all u of H. Then, if $u_1 = u + v$,

$$G[u_1, s] - G[u, s] = \sum_{i=1}^{\infty}(m) (z_{i1} - z_i) \int_{(0,1)} L[u + rv; p_i, s] dr. \qquad [u_1 = (z_{11}, z_{21}, \cdots)]$$

Since L is linear, we know that

$$L[u + rv; v_n, s] = \sum_{i=1}^{n} z_i' L[u + rv; p_i, s], \qquad (z_i' = z_{i1} - z_i)$$

where $v_n = z_1' p_1 + \cdots + z_n' p_n$. Moreover,

$$L[u + rv; v, s] - L[u + rv; v_n, s] = L[u + rv; v - v_n, s]. \tag{29}$$

* Paper I, Theorem IX.

Consequently, for a fixed value of r, the modulus of the left member of (29) is at most

$$KM(r - r_n), \tag{30}$$

and hence approaches zero for $n = \infty$, uniformly for $0 \leq r \leq 1$.

In view of (28) it follows that Corollary (1) will be established if we show that

$$\lim_{n=\infty} \int_E \left[\int_{(0,1)} (L[u + rv; r, s] - L[u + rv; r_n, s])\, dr \right]^2 ds = 0. \tag{31}$$

By an application of the Schwarz inequality to the inner integral in (31), it is seen that, for a given value of n, the integral in (31) is at most

$$\int_E \left[\int_{(0,1)} L^2[u + rv; r - r_n, s]\, dr \right] ds \tag{32}$$
$$= \int_{(0,1)} dr \int_E L^2[u + rv; r - r_n, s]\, ds \leq K^2(d - c)\, M^2(r - r_n).$$

The interchange of integrations made in obtaining (32) was permissible* under our present hypotheses. The inequality in (32) is a result of (30). From (32) it is evident that (31) is true.

3. The Fourier coefficients of $G[u, s]$. Let $(y_i(s); i = 1, 2, \cdots)$ be a system of functions, unitary and orthogonal on E and complete for the class W. For a given function u of H, $G[u, s]$ will possess Fourier coefficients $(F_i[u]; i = 1, 2, \cdots)$, where

$$F_i[u] = \int_E G[u, s]\, y_i(s)\, ds, \tag{33}$$

which satisfy the equation

$$\int_E G^2[u, s]\, ds = \sum_{i=1}^{\infty} F_i^2[u].$$

It is obvious that, if $G[u, s]$ is a linear functional, then F_i is linear in the sense of Part I. Further properties of the (F_i) are considered in the next theorem.

THEOREM 4. *The functionals* (F_i) *are continuous in the sense of Part I. Furthermore, if* $G[u, s]$ *possesses a differential at* u_0 *in* H, *the functionals* (F_i) *possess differentials at* u_0 *in the sense of Part I.*

As a consequence of (33), it is seen that

$$\int_E (G[u_n, s] - G[u, s])^2\, ds = \sum_{i=1}^{\infty} (F_i[u_n] - F_i[u])^2. \tag{34}$$

* Cf. de la Vallée Poussin, *Intégrales de Lebesgue*, p. 53.

If a sequence (u_n) satisfies $\underset{n=\infty}{\text{l. m.}}\, u_n = u$, the whole expression (34), and, therefore, every term in the infinite series, approaches zero for $n = \infty$. This establishes the continuity of the (F_i).

If G possesses a differential at u_0, it is seen from equation (26) that

$$(35) \qquad F_i[u] - F_i[u_0] = L_i[u_0; r] + Mr \int_E e(u_0, r, s)\, g_i(s)\, ds,$$

where L_i is the ith Fourier coefficient of L. The equation (35) shows that L_i is the differential of F_i provided that the integral entering in (35) approaches zero with Mr. This fact is easily established by use of the properties of $e(u_0, r, s)$, because

$$\left[\int_E e(u_0, r, s)\, g_i(s)\, ds\right]^2 \leqq \left(\int_E g_i^2(s)\, ds\right) \int_E e^2(u_0, r, s)\, ds.$$

This completes the proof of Theorem 4.

4. **The functions in Hilbert space corresponding to** $G[u, s]$. Let us define a function $g(\zeta, s)$, in which ζ is in Hilbert space, by the equation $g(\zeta, s) = G[u, s]$, where $u = \zeta$. We shall say that $g(\zeta, s)$ corresponds to $G[u, s]$ and shall denote this fact by the notation $g(\zeta, s) = G[u, s]$. With the aid of equations (3) and (4) it is verified that, if $\underset{n=\infty}{\text{l. m.}}\, \zeta_n = \zeta$, then $\underset{n=\infty}{\text{l. m.}}\, g(\zeta_n, s) = g(\zeta, s)$. If $L[u, s]$ is linear, its corresponding function $l[\zeta, s]$ is given by the right member of (24) and hence is a linear form in the variables $(z_1, z_2, \cdots)$ converging in the mean.

THEOREM 5. *At the point u in H suppose that $G[u, s]$ possesses the differential $L[u; r, s]$, and let $g(\zeta, s) = G[u, s]$. Then, at the point $\zeta = u$, there exist partial pseudo-derivatives $g_{z_i}(\zeta, s)$ given by the equations*

$$g_{z_i}(\zeta, s) = l(\zeta; p_i, s) \qquad (l(\zeta; p_i, s) = L[u; p_i, s];\ i = 1, 2, \cdots).$$

For simplicity consider only the case $i = 1$. Let $u_1 = u + hp_1$. With the aid of (26) we obtain

$$\frac{\Delta G}{h} = \frac{g(z_1 + h, z_2, z_3, \cdots, s) - g(\zeta, s)}{h} = \frac{G[u_1, s] - G[u, s]}{h}$$
$$= L[u; p_1, s] + e(u, hp_1, s),$$

where use has been made of the fact that $Mp_1 = 1$. As a consequence of the properties of $e(u, r, s)$ it is seen that

$$\lim_{h=0} \int_E \left(\frac{\Delta g}{h} - L[u; p_1, s]\right)^2 ds = 0,$$

which establishes the theorem.

5. An implicit functional equation. Consider the functional equation

$$G[u; s, t] = 0, \tag{36}$$

where, for every t on an interval $|t - t_0| \leq k$, the functional $G[u; s, t]$ is of the type considered previously in this part of the paper. Suppose that $u = u_0, t = t_0$ satisfies (36). Then, we shall seek to determine a function $u(x, t)$ which, for every t, is a function of H and which has the property that $G[u(*, t); s, t]$, when t is fixed, is zero almost everywhere on E. The argument x in $u(x, t)$ was replaced here by "*" in order to emphasize the character of the dependence of G on u. The Riesz-Fischer theorem shows that the problem proposed is equivalent to the problem solved in Theorem 5, Part I, for the system

$$F_i[u, t] = 0, \qquad (i = 1, 2, \cdots; F_i[u, t] = \int_E G[u; s, t]\, y_i(s)\, ds) \tag{37}$$

In regard to (36) we shall make the following assumptions:

(a) For every t, $G[u; s, t]$ possesses a differential $L[u; r, s, t]$ at all points u in H.

(b) If $\operatorname*{l.m.}_{n=\infty} u_n = u$, $\operatorname*{l.m.}_{n=\infty} r_n = r$, and $\lim_{n=\infty} t_n = t$, then

$$\operatorname*{l.m.}_{n=\infty} G[u_n; s, t_n] = G[u; s, t],$$
$$\operatorname*{l.m.}_{n=\infty} L[u_n; r_n, s, t_n] = L[u; r, s, t].$$

(c) The series (20), formed for the system (37), converge uniformly for u in H and for $|t - t_0| \leq k$.

From (a), with the assistance of Theorem 4, it it seen that F_i in (37) possesses the differential

$$L_i[u; r, t] = \int_E L[u; r, s, t]\, y_i(s)\, ds.$$

With the aid of (b) an application of the Schwarz inequality shows that

$$\begin{aligned}(F_i[u_n, t_n] - F_i[u, t])^2 &= \left[\int_E (G[u_n; s, t_n] - G[u; s, t])\, y_i(s)\, ds\right]^2, \\ &\leq \int_E (G[u_n; s, t_n] - G[u; s, t])^2 ds,\end{aligned}$$

which approaches zero for $n = \infty$. Hence, F_i has the continuity property postulated for system (18) of Part I. It may be shown in a similar fashion that the differential L_i of the present section has the continuity assumed for the L_i of that system.

In (c) it should be remarked that the first series of (20) always converges for a system (37) since the F_i are the Fourier coefficients of G. The force of the condition (c) lies in the assumption of uniformity for this convergence. The uniform convergence of the third series of (20) for system (37) would be insured if we should postulate uniform convergence for the series

$$\sum_{i=1}^{\infty} \int_E (L[u, p_i, s, t] - y_i(s))^2 ds.$$

The consequences we have just derived from the assumptions (a), (b), and (c) permit us to state the next theorem without further proof.

THEOREM 6. *Suppose that* (36) *satisfies* (a), (b), *and* (c), *and that the infinite determinant* (21), *formed for the system* (37) *is different from zero. Then, the hypotheses of Theorem 5, Part 1, are satisfied by* (37), *and the concluding statement of that theorem in regard to the existence of a solution $u(x,t)$ applies to* (36).

6. **A pseudo-differential equation.** For every value of t on $|t - t_0| < k$ suppose that the function $u(x,t)$ is a function of H on the interval (a,b). Consider the equation

$$u_t(x,t) = G[u(*,t);t,x], \tag{38}$$

where for every t, $G[u;t,x]$ is a functional operation of the type defined in § 1, with the interval E replaced by the interval (a,b). In (38) u_t represents the pseudo-derivative of u with respect to t, and it should be remembered that, by the definition of the pseudo-derivative, $u_t(x,t)$ is required to be a function of H for every value of t. The "$*$" is used in (38) to emphasize the fact that G is a functional of u.

Let P represent the class of all functions $u(x,t)$ which possess pseudo-derivatives $u_t(x,t)$ for some values of t. Then, we wish to state hypotheses under which (38) will have a solution $u(x,t)$ belonging to P and satisfying a given initial condition $u(x,0) = u_0(x)$, where $u_0(x)$ is in H.

Before discussing (38) let us recall certain facts about pseudo-derivatives previously established by the author*. Let $\zeta(t) = [z_1(t), z_2(t), \cdots]$ and $\eta(t) = [k_1(t), k_2(t), \cdots]$ belong to Hilbert space for every t, and let u and w be functions of (x,t) satisfying the relations $\zeta(t) \sim u(x,t)$, $\eta(t) \sim w(x,t)$. If $w(x,t) = u_t(x,t)$ then $dz_i/dt = k_i$ and, moreover,

$$\lim_{h=0} \sum_{i=1}^{\infty} \left(\frac{z_i(t+h) - z_i(t)}{h} - k_i(t) \right)^2 = 0. \tag{39}$$

* Paper I, Theorems VII and VIII.

for every value of t. Conversely, if $\zeta(t)$ and $\eta(t)$ satisfy (39), it follows that the pseudo-derivative u_t exists and is defined by the equation $u_t(x,t) = w(x,t)$. In view of these statements it is seen that, if $u(x,t)$ is in the class P and if $\zeta(t) = u(x,t)$, then $d\zeta(t)/dt = u_t(x,t)$, where

$$\frac{d\zeta(t)}{dt} = \left(\frac{dz_1(t)}{dt}, \frac{dz_2(t)}{dt}, \cdots\right).$$

For every t and u of H define the functional $G_i[u,t]$ by the equation

$$\begin{aligned} G_i[u,t] &= \int_{(ab)} G[u;\, t, x]\, p_i(x)\, dx, \\ &= g_i(\zeta, t). \qquad (u = \zeta;\ g_i(\zeta, t) = G_i[u,t]) \end{aligned}$$

It is seen that, if (38) is satisfied by a function $u(x,t)$ of P, the point $\zeta(t) = u(x,t)$ satisfies the system of equations

$$\text{(40)} \qquad \frac{dz_i}{dt} = g_i(\zeta, t). \qquad (i = 1, 2, \cdots)$$

Conversely, if $\zeta = \zeta(t)$ is a solution of (40) whose coordinates $z_i(t)$ satisfy (39) with k_i replaced by dz_i/dt, it follows that the function $u(x,t)$ corresponding to $\zeta(t)$ is in the class P and satisfies (38). A system (40) whose solution has the desired property (39) has been treated previously by the author.* Sufficient conditions will now be imposed on (38) so that the system (40) corresponding to it can be treated by the theorem referred to.

Assume in the future that $G[u;\, t, x]$ satisfies the following conditions:

(a) If $\underset{n=\infty}{\text{l. m.}}\, u_n(x) = u(x)$ and $\lim_{n=\infty} t_n = t$, then

$$\underset{n=\infty}{\text{l. m.}}\, G[u_n;\, t_n, x] = G[u;\, t, x].$$

(b) There exist positive constants A and B such that

$$M[G[u_1;\, t, x] - G[u_2;\, t, x]] < A\, M(u_1 - u_2) + B\, |t_1 - t_2|,$$

for all points (t_1, t_2) and all functions (u_1, u_2) of H.

(c) Corresponding to the function $u_0(x)$ in H there exists an upper bound $D > 0$ for the quantity $MG[u_0;\, t, x]$ for $|t - t_0| \leq k$.

THEOREM 7. *If (a), (b), and (c) are satisfied, then, for $|t - t_0|$ sufficiently small there exists uniquely, among functions of the class P, a function $u(x,t)$ which satisfies (38) and, also, the condition $u(x, 0) = u_0(x)$.*

* Paper II, Theorem IV.

Consider the system (40) corresponding to (38). As a consequence of (a) it is seen that, if $\underset{n=\infty}{\text{l.m.}}\, \zeta_n = \zeta$ and $\lim\limits_{n=\infty} t_n = t$, then, if

$$(\zeta_n = u_n,\ \zeta = u)$$

$$0 = \lim_{n=\infty} \int_{(ab)} (G[u_n; t_n, x] - G[u; t, x])^2\, dx = \lim_{n=\infty} \sum_{i=1}^{\infty} [g_i(\zeta_n, t_n) - g_i(\zeta, t)]^2.$$

Therefore, each function $g_i(\zeta, t)$ possesses continuity, in the weak sense with respect to ζ and in the ordinary sense in t, simultaneously in its arguments. From (b) it follows that

$$\sum_{i=1}^{\infty} [g_i^2(\zeta_1, t_1) - g_i^2(\zeta_2, t_2)]^2 \leq (AM(\zeta_1 - \zeta_2) + B|t_1 - t_2|)^2,$$

for all points (ζ_1, ζ_2) in Hilbert space and for all (t_1, t_2) on $|t - t_0| \leq k$. As an obvious consequence of (c) it is seen that D^2 is an upper bound for

$$\sum_{i=1}^{\infty} g_i^2(\zeta_0, t). \qquad (\zeta_0 = u_0;\ |t - t_0| \leq k)$$

The consequences of (a), (b) and (c) just established constitute the assumptions under which we may state* that, for $|t - t_0|$ sufficiently small, (40) has a unique solution $\zeta = \zeta(t)$ $(\zeta(0) = \zeta_0)$, in Hilbert space, which satisfies (39) with $k_i = dz_i/dt$. From our previous discussion of the relationship between (38) and (40) it is seen that Theorem (7) has been completely established.

Various generalizations of the results of Parts I and II are immediately obvious. Instead of taking H as the class of all integrable functions we could have restricted it to those functions u satisfying $M(u - u_0) \leq h$, where u is a particular integrable function, and where h is a given positive constant. Obvious changes would permit all proofs given to extend to this more general case. The variables (s, x, t) used above to represent single variables could be thought of as being m-partite variables and the necessary modifications of the hypotheses in the work above are easily determined.

* Paper II, Theorem IV and Corollary (2) to Theorem IV.

UNIVERSITY OF MINNESOTA,
March 20, 1921.

PERIODICALLY CLOSED CHAINS OF REDUCED FRACTIONS.

By A. Arwin.

Introduction. In § 1 we first have to prove the existence of the positively and in § 5 of the negatively reduced fractions. With integers P_ν, Q_ν and D, we call the element $(P_\nu + \sqrt{D})/Q_\nu$ a positively reduced fraction, if the inequalities $0 < (P_\nu - \sqrt{D})/Q_\nu < 1$, $(P_\nu + \sqrt{D})/Q_\nu > 1$ are satisfied, and the element $(\sqrt{D} + P_\tau)/Q_\tau$ a negatively reduced fraction, if $0 < (\sqrt{D} - P_\tau)/Q_\tau < 1$, $(\sqrt{D} + P_\tau)/Q_\tau > 1$. After the positively reduced fractions in § 1 have been defined, their chain formations will be studied, that is the formation

$$(1) \qquad \frac{P_\rho + \sqrt{D}}{Q_\rho} = \beta_\rho - \frac{P_{\rho+1} - \sqrt{D}}{Q_\rho} = \beta_\rho - \cfrac{1}{\cfrac{P_{\rho+1} + \sqrt{D}}{Q_{\rho+1}}}, \text{ \&c.}$$

where β_ρ is always the integer next over the greatest integer in $(P_\rho + \sqrt{D})/Q_\rho$, and then we have to prove our fundamental theorem: *Whichever element opens a chain of the nature* (1), *we always end in a periodically closed chain of reduced fractions, and only a finite number of such chains exists.* Next we shall see that these chains include all possible positively reduced fractions. Moreover, the chain complementary to that derived from $(P_\nu + \sqrt{D})/Q_\nu$ as its first element being by definition the chain derived from the first element $(P_\nu + \sqrt{D})/Q_{\nu-1}$ where $P_\nu^2 - D = Q_{\nu-1}Q_\nu$, it will be shown that the complementary chain is periodic and the inverse of the original chain. These two chains may be different, or they may coincide. In the latter case we shall prove that the continued fraction is symmetrical. The different types of symmetry will be given, and the reason for their symmetry will be shown. At the end of § 1 another fundamental formula, valid for all formations of fractions, will be proved, and it will also be shown how the closed chains give solutions to Pell's equation. In § 2 attention is called to some types of remarkably regular formations of chains, so that we may survey them better. In § 3 we prove the following theorem: The necessary and sufficient condition that the indeterminate equation $x^2 - Dy^2 = A$ have a solution is that the chain formation (1) shall run into the special, periodic chain that contains

in it $Q_0 = 1$, when all solutions may easily be taken from the related continued fraction. At the same time we discuss the problem of writing a given number A in the form (a, b, c), that is of solving in integers x and y the equation $ax^2 + 2bxy + cy^2 = A$. In § 4 the chain containing the element $Q_\nu = 1$ or $(P_\nu + \sqrt{D})/1$, the so called principal chain which customarily is used in solving the Pellian equation is investigated. In § 5 the negatively reduced fractions are introduced, their existence proved, also their connection with the positively reduced fractions. Then, proceeding from the results found, we shall see that Gauss's classical theory of binary quadratic forms, al least in its main characteristics, can be given by our theory of periodically closed chains of fractions, and in particular that the number of the positively closed chains is most intimately connected with the number of classes for a given determinant D, of quadratic forms and the negatively with the number of ideal classes in the domain $K(\sqrt{D})$. Finally in § 6 it will briefly be shown how quadratic forms with negative D can also be brought under this generalized theory of continued fractions.

§ 1.

We shall now prove the existence of positively reduced fractions $(P_\nu + \sqrt{D})/Q_\nu$, defined by the two inequalities $0 < (P_\nu - \sqrt{D})/Q_\nu < 1$, $(P_\nu + \sqrt{D})/Q_\nu > 1$.

THEOREM 1: *The number of the periodically closed chains is finite.* We form the chain of values

$$(2') \qquad P_0^2 - D = Q_1 Q_0,\ P_1^2 - D = Q_2 Q_1, \text{ \&c.}$$

with $P_\tau + P'_{\tau+1} \equiv 0 \bmod Q_\tau$ and determine for $0 < (P'_{\nu+1} - \sqrt{D})/Q_\nu > 1$ the number n from $0 < (P'_{\nu+1} - nQ_\nu - \sqrt{D})/Q_\nu < 1$ to be $n = [(P'_{\nu+1} - \sqrt{D})/Q_\nu]$, where [] as usual is the greatest integer in $(P'_{\nu+1} - \sqrt{D})/Q_\nu$, and $P'_{\nu+1} - nQ_\nu = P_{\nu+1}$. Obviously we must have $P_{\nu+1} < Q_\nu$ for $Q_\nu > D$. For, $P_{\nu+1} - Q_\nu - \sqrt{D} > 0$, since $(P_{\nu+1} - Q_\nu)^2 - D = Q_\nu Q_{\nu+1} > 0$; and hence the inequalities $0 < (P_{\nu+1} - \sqrt{D})/Q_\nu < 1$ and $0 < (P_{\nu+1} - Q_\nu - \sqrt{D})/Q_\nu < 1$ are incompatible. If $P'^2_{\nu+1} - D < 0$ we take n from the inequality $0 < (P'_{\nu+1} + nQ_\nu - \sqrt{D})/Q_\nu$ and get $n = [(\sqrt{D} - P'_{\nu+1})/Q_\nu] + 1$. Therefore we may also write (2') as the following chain of fractions

$$(2'') \qquad \begin{aligned} (P_0 + \sqrt{D})/Q_0 &= \beta_0 - (P_1 - \sqrt{D})/Q_0 = \beta_0 - \frac{1}{(P_1 + \sqrt{D})/Q_1}, \\ (P_1 + \sqrt{D})/Q_1 &= \beta_1 - (P_2 - \sqrt{D})/Q_1 = \beta_1 - \frac{1}{(P_2 + \sqrt{D})/Q_2}. \end{aligned}$$

in which, since $0 < (P_{\nu+1} - \sqrt{D})/Q_\nu < 1$, β_ν is the integer next above the greatest integer in $(P_\nu + \sqrt{D})/Q_\nu$. The chain formation (2″) is preferable to (2′). In the sequel we may therefore refer to the two chains (2′) and (2″) as both uniquely determined and identical with each other. Let us now suppose $Q_\tau > D$. Then we get $P_{\tau+1}^2 - D = Q_\tau \cdot Q_{\tau+1} < P_{\tau+1}^2 < Q_\tau^2$, hence $Q_{\tau+1} < Q_\tau$ and thus as far as $Q_\tau > D$ the set of values $Q_\nu (\nu > \tau)$ succesively decreases until we get a $Q_\rho < D$ with a $P'_{\rho+1}$, that is taken as least positive remainder from $-P_\rho \equiv P'_{\rho+1}$ mod. Q_ρ giving $P'^2_{\rho+1} - D < 0$. If in this case we take n as above, we get the following inequalities $D + \sqrt{D} > P'_{\rho+1} + nQ_\rho > \sqrt{D}$ because of the inequality $n = (\sqrt{D} - P'_{\rho+1})/Q_\rho - \mu + 1 < (D + \sqrt{D} - P'_{\rho+1})/Q_\rho$, which in its turn is always satisfied for values $\mu < 1$, $Q_\rho < D$. With regard to $(P'_{\rho+1} + nQ_\rho)^2 - D < (D + \sqrt{D})^2 - D = D(D + 2\sqrt{D}) = M(D)$ we must in $(P'_{\rho+1} + nQ_\rho)^2 - D = P_{\rho+1}^2 - D = Q_\rho Q_{\rho+1}$ have $Q_{\rho+1} < M(D)$. When $Q_{\rho+1} > D$ the $Q_\tau (\tau > \varrho + 1)$, as we have shown, may be reduced so as to be less than D. In the cases $Q_{\rho+1} < D$, $P'^2_{\rho+2} - D < 0$ the above process upon repetition gives $Q_{\rho+2} < M(D)$. We now have $P_{\rho+1} = P'_{\rho+1} + nQ_\rho < D + \sqrt{D} < M(D)$ and when $Q_{\rho+1} > D$ also $P_{\rho+2} < Q_{\rho+1} < M(D)$. Hence in every case the set of values P_ν, Q_ν in (2″) can be reduced permanently below the limit $M(D)$. As however we may form fractions indefinitely, we must in every chain get the same set of fractions periodically repeated and must also have the number of existing chains finite, and thus our first fundamental theorem is proved. We shall now prove

THEOREM 2. *The fractions in the periodically closed chains are positively reduced.*

We have first to prove that, given the chain

$$
(2') \qquad
\begin{gathered}
\cdots\cdots\cdots \\
P_{\nu-1}^2 - D = Q_{\nu-2} Q_{\nu-1}, \\
P_\nu^2 - D = Q_{\nu-1} Q_\nu, \\
P_{\nu+1}^2 - D = Q_\nu Q_{\nu+1}, \\
\cdots\cdots\cdots
\end{gathered}
$$

and for example $0 < (P_\nu - \sqrt{D})/Q_\nu < 1$, $(P_\nu + \sqrt{D})/Q_\nu > 1$ all subsequent fractions satisfy precisely similar inequalities. From $(P_\nu + \sqrt{D})/Q_\nu = \beta_\nu - (P_{\nu+1} - \sqrt{D})/Q_\nu > 1$ we conclude $\beta_\nu \geq 2$. Hence $(P_{\nu+1} + \sqrt{D})/Q_\nu > 1$. Then $[(P_{\nu+1} + \sqrt{D})/Q_\nu][(P_{\nu+1} - \sqrt{D})/Q_{\nu+1}] = 1$, $[(P_{\nu-1} - \sqrt{D})/Q_\nu][(P_{\nu-1} + \sqrt{D})/Q_{\nu-1}] = 1$ will directly give us $0 < (P_{\nu+1} - \sqrt{D})/Q_{\nu+1} < 1$, $(P_{\nu+1} + \sqrt{D})/Q_{\nu+1} > 1$ and thus similar inequalities are valid for all fractions later in the chain. Further, from the law of formation in the chain, all elements

$(P_{\tau+1} - \sqrt{D})/Q_\tau$ from the second column of (2″) lie in the interval between 0 and 1. If therefore the element $(P_\nu + \sqrt{D})/Q_\nu$ is the opening element of a periodically closed chain in the first column, the conjugate of the complementary element $(P_\nu + \sqrt{D})/Q_{\nu-1}$ must be found in the second column of the last row in order to close the chain. From $P_\nu^2 - D = -Q_{\nu-1}Q_\nu > 0$, $[(P_\nu - \sqrt{D})/Q_{\nu-1}][(P_\nu + \sqrt{D})/Q_\nu] = 1$ we have $(P_\nu + \sqrt{D})/Q_\nu > 1$, and similarly for each element in the periodically closed chain. Now for any element $(P_\tau - \sqrt{D})/Q_\tau > 0$ having $(P_\tau - \sqrt{D})/Q_\tau < 1$ we have proved that all elements in the periodic chain must satisfy the same inequality. Let us therefore assume $(P_\tau - \sqrt{D})/Q_\tau > 1$ for all τ. Then we must also for all τ have $Q_{\tau-1} > Q_\tau$. For if there were only one $\tau = \chi$ with $Q_{\chi-1} < Q_\chi$ we should have $P_\chi^2 - D = Q_{\chi-1}Q_\chi < Q_\chi^2$, $[(P_\chi - \sqrt{D})/Q_\chi][(P_\chi + \sqrt{D})/Q_\chi] < 1$, contrary to $(P_\chi + \sqrt{D})/Q_\chi > 1$, $(P_\chi - \sqrt{D})/Q_\chi > 1$. But since for all τ in the complementary chain $P_\rho^2 - D = Q_\rho Q_{\rho-1}$, &c. we have $Q_{\tau-1} > Q_\tau$, it would then follow that the chain could not be closed. Thus Theorem 2 is proved.

Let us assume the chain to be closed first with this row of elements

$$(3) \qquad \frac{P_{\nu+\mu} + \sqrt{D}}{Q_{\nu+\mu}} = \beta_{\nu+\mu} - \frac{P_{\nu+\mu+1} - \sqrt{D}}{Q_{\nu+\mu}} = \beta_{\nu+\mu} - \frac{1}{(P_{\nu+\mu+1} + \sqrt{D})/Q_{\nu+\mu+1}}.$$

The chain of the element $(P_{\nu+\mu} + \sqrt{D})/Q_{\nu+\mu}$ and also that of the element $(P_{\nu+\mu+1} + \sqrt{D})/Q_{\nu+\mu}$ is uniquely determined by the stated inequalities. If the periodic chain were now to be opened with the row in (3), the formation of chain with $(P_{\nu+\mu+1} + \sqrt{D})/Q_{\nu+\mu}$ might then not bring us out of the closed sets of values from the periodic part of the chain, but according to the given inequalities, assumed valid from the element $(P_\nu + \sqrt{D})/Q_\nu$, this same chain formation must restore the element $(P_\nu - \sqrt{D})/Q_\nu$. Hence $(P_\nu + \sqrt{D})/Q_\nu$ itself opens the chain. As a corollary we have immediately: Where the positively reduced element $(P_\nu + \sqrt{D})/Q_\nu$ forms a continued fraction, the complementary element $(P_\nu + \sqrt{D})/Q_{\nu-1}$ will give the inverse continued fraction.

We shall next prove the important theorem:

THEOREM 4. *The reduced element* $(P_\tau + \sqrt{D})/Q_\tau$ *satisfies the inequalities*

$$P_\tau \leq D, \qquad Q_\tau < D.$$

Having $-P_\tau \equiv \alpha_\tau \bmod. Q_\tau$, $\alpha_\tau^2 - D < 0$ we always obtain $Q_\tau < D$ from $D - \alpha_\tau^2 = Q_\tau M$. Now, having $P_{\tau+1} < Q_\tau$ and thus $Q_\tau > Q_{\tau+1}$ with $P_{\tau+1}^2 - D > 0$, we obtain from $(P_{\tau+1} + \sqrt{D})/Q_\tau > 1$ $(P_{\tau+1} + \sqrt{D})/Q_\tau = \mu_\tau + (\sqrt{D}$

$-\beta_\tau)/Q_\tau$, $\mu_\tau > 1$ as the greatest integer in $(P_{\tau+1}+\sqrt{D})/Q_\tau$ and $\beta_\tau > 0$, for since $\beta_\tau = -\gamma_\tau$, $\gamma_\tau > 0$ the relation $P_{\tau+1} - \gamma_\tau = \mu_\tau Q_\tau$ is not possible with $P_{\tau+1} < Q_\tau$. From $\sqrt{D} - \beta_\tau > 0$ and $D - \beta_\tau^2 = Q_\tau N$ we once more conclude $Q_\tau < D$ for all these ι. It therefore remains only to examine such cases as $Q_{\tau+1} > Q_\tau$, which follows from $P_{\tau+1} = \alpha_\tau + nQ_\tau$, as well as the case $P_{\tau+1} > Q_\tau$, $Q_\tau > Q_{\tau+1}$. When $P^2_{\tau+1} - D = Q_\tau Q_{\tau+1}$ we have just seen that $Q_\tau < D$, and either

$$-P_{\tau+1} \equiv \alpha_{\tau+1} \bmod. Q_{\tau+1}, \qquad \alpha^2_{\tau+1} - D < 0,$$

whence $Q_{\tau+1} < D$, or

$$-P_{\tau+1} \equiv P_{\tau+2} \bmod. Q_{\tau+1}, \qquad P^2_{\tau+2} - D > 0.$$

In this, since $\beta_{\tau+1} \geq 2$, we must have $P_{\tau+1} + P_{\tau+2} > 2Q_{\tau+1}$, $P_{\tau+1} > Q_{\tau+1} - P_{\tau+2}$ and get as before $Q_{\tau+1} < D$. Since now in all cases $Q_\tau \leq D$, $Q_{\tau+1} \leq D$, we infer from $P^2_{\tau+1} - D < D \cdot D$ that $P^2_{\tau+1} < D^2$, and hence that $P_\rho < D$, $Q_\rho \leq D$, for all values of ρ, which establishes theorem 4.

As example of periodically closed chain we give

$$\frac{11+\sqrt{79}}{14} = 2 - \frac{17-\sqrt{79}}{14} = 2 - \cfrac{1}{\cfrac{17+\sqrt{79}}{15}}.$$

$$\frac{11+\sqrt{79}}{7} = 3 - \frac{10-\sqrt{79}}{7} = 3 - \cfrac{1}{\cfrac{10+\sqrt{79}}{3}}.$$

$$\frac{17+\sqrt{79}}{15} = 2 - \frac{13-\sqrt{79}}{15} = 2 - \cfrac{1}{\cfrac{13+\sqrt{79}}{6}}.$$

$$\frac{10+\sqrt{79}}{3} = 7 - \frac{11-\sqrt{79}}{3} = 7 - \cfrac{1}{\cfrac{11+\sqrt{79}}{14}}.$$

$$\frac{13+\sqrt{79}}{6} = 4 - \frac{11-\sqrt{79}}{6} = 4 - \cfrac{1}{\cfrac{11+\sqrt{79}}{7}}.$$

One of the following cases is to be expected: Either the two chains, the original and the complementary, are distinct or coincident. We now discuss the latter case. Given that $(P_{\tau+1}+\sqrt{D})/Q_{\tau+1}$ follows $(P_\tau+\sqrt{D})/Q_\tau$, we ask in what order are the elements $(P_\tau+\sqrt{D})/Q_{\tau+1}$, $(P_{\tau+1}+\sqrt{D})/Q_\tau$? Obviously the elements in the complementary chain are to be taken as the conjugates

of the elements in the second column and read from below. In the original chain the order therefore must be $(P_{\tau-1}+\sqrt{D})/Q_{\tau}$ to $(P_{\tau}+\sqrt{D})/Q_{\tau-1}$ from above. By simple transposition of the last row of elements to the initial part of the periodic chain we can then always arrange these chains to be symmetrical. About the middle of any chain the elements $(P_{\rho}+\sqrt{D})/Q_{\rho}$, $(P_{\rho}+\sqrt{D})/Q_{\rho-1}$ must either be successive or situated with specific properties on opposite sides of the element $(P_{\rho-1}+\sqrt{D})/Q_{\rho-1}$, according to the type of symmetry involved. In the former case the "middle element" is characterized by the relation $P_{\nu}=P_{\nu-1}$. As, further, we have throughout $P_{\rho}+P_{\rho-1}=\beta_{\rho}Q_{\rho}$ and for this middle element $P_{\nu-1}=P_{\nu}$, there results in this case the important equality

$$2P_{\nu}=\beta_{\nu}Q_{\nu}. \tag{5}$$

In the latter case we have

$$\begin{array}{ll} & \cdots\cdots\cdots\cdots\cdots\cdots \\ \downarrow & P_{\rho}^2-D=Q_{\rho-1}Q_{\rho}=Q_{\rho-1}Q_{\rho}, \\ (6) & P_{\rho-1}^2-D=Q_{\rho}Q_{\rho-1}=Q_{\rho}Q_{\rho}, \\ & P_{\rho}^2-D=Q_{\rho-1}Q_{\rho-2}=Q_{\rho}Q_{\rho-1}, \uparrow \\ & \cdots\cdots\cdots\cdots\cdots\cdots \end{array}$$

and the middle element characterized by $Q_{\nu-1}=Q_{\nu}$. Hence

$$D=P_{\nu-1}^2-Q_{\nu-1}^2. \tag{7}$$

Calling the middle element in the former case an element of the first category and in the latter an element of the second category we may formulate the following theorem.

THEOREM 5. *The periodic chain being opened with an element of the second category ($D=P_0^2-Q_0^2$) and having $Q_1 \neq Q_0$, the continued fraction formed will be symmetric throughout, otherwise the symmetry will only appear after β_0, the first coefficient in the continued fraction.*

Arranging the chain from the middle element as an initial element we find that the initial element, now the middle element, must have just the same properties as any middle element, and therefore in a symmetrical chain there must exist such two specific elements belonging to one of the categories mentioned. And further, in the first case the initial element of the first column and the last element of the second will be conjugate and therefore after the first element the complementary and the original coincide, and the symmetry is complete. In the latter case the element in the second row of the original chain and the last element from the second column are the first to be conjugate, giving symmetry only after β_0.

THEOREM 6. *If in the symmetrical chain one element of each category exists the period of the continued fraction formed is odd, if there are two elements of the same category the period is even.*

If the initial element and also the middle element are of the second category, the symmetry is complete and in the middle of the chain we find the elements $(P_\nu + \sqrt{D})/Q_\nu$ and $(P_{\nu+1} + \sqrt{D})/Q_\nu$ with the same denominators following each other. Hence the period must be even. And if the initial and also the middle element are of the first category, elements with the same denominators will group themselves symmetrically around $(P_{\nu+1} + \sqrt{D})/Q_{\nu+1}$, but as in this case symmetry appears only after β_0, the period must be even. In the same manner we have for symmetrical chains containing one element of either category an odd period, and the theorem is proved. Thus we may for the different cases set up the following scheme.

Period $= 2\nu$, $Q_1 \neq Q_0$ and $Q_{2\nu-1} \neq Q_{2\nu}$,

$$Q_{\nu-\tau} = Q_{\nu+\tau}, \qquad (\tau = 0, 1 \cdots \nu)$$

Period $= 2\nu$, $Q_1 \neq Q_0$ and $Q_{2\nu-1} = Q_{2\nu}$, $D = P^2_{2\nu} - Q^2_{2\nu} = P^2_\nu - Q^2_\nu$,

$$Q_{\nu-\tau-1} = Q_{\nu+\tau}, \qquad (\tau = 0, 1 \cdots \nu - 1)$$

Period $= 2\nu + 1$, $Q_1 \neq Q_0$ and $Q_{2\nu} \neq Q_{2\nu-1}$, $D = P^2_{\nu+1} - Q^2_{\nu-1}$,

$$Q_{\nu-\tau} = Q_{\nu+\tau+1}, \qquad (\tau = 0, 1 \cdots \nu)$$

Period $= 2\nu + 1$, $Q_1 \neq Q_0$ and $Q_{2\nu} = Q_{2\nu-1}$,

$$Q_{\nu-\tau} = Q_{\nu+\tau}, \qquad (\tau = 0, 1 \cdots \nu)$$

Assuming now that we have an arbitrarily orientated, not necessarily closed, chain with the initial element $\chi_0 = (P_0 + \sqrt{D})/Q_0$ and $\chi_\tau = (P_\tau + \sqrt{D})/Q_\tau$ standing anywhere in the chain, we may then formulate the following theorem.

THEOREM 7. *Having*

$$\frac{y_{\tau-1}}{z_{\tau-1}} = [\beta_0, \beta_1, \cdots \beta_{\tau-1}]$$

we may always take from the chain the relation

$$(Q_0 y_{\tau-1} - P_0 z_{\tau-1})^2 - D z^2_{2-1} = Q_0 Q_\tau.$$

Because*

$$(8') \qquad \chi_0 = \frac{\chi_\tau y_{\tau-1} - y_{\tau-2}}{\chi_\tau z_{\tau-1} - z_{\tau-2}}, \text{ and } \frac{y_{\tau-2}}{z_{\tau-2}} = [\beta_0, \beta_1, \cdots \beta_{\tau-2}],$$

we get

$$[(P_0 + \sqrt{D})/Q_0]\,[(P_\tau + \sqrt{D})/Q_\tau]\, z_{\tau-1} - [(P_0 + \sqrt{D})/Q_0]\, z_{\tau-2}$$
$$= [(P_\tau + \sqrt{D})/Q_\tau]\, y_{\tau-1} - y_{\tau-2},$$

giving the two equations

$$(8'') \qquad \begin{aligned} (P_0 P_\tau + D)\, z_{\tau-1} - P_\tau Q_0 y_{\tau-1} - P_0 Q_\tau z_{\tau-2} &= - Q_0 Q_\tau y_{\tau-2}, \\ (P_0 + P_\tau)\, z_{\tau-1} - Q_0 y_{\tau-1} &= Q_\tau z_{\tau-2}. \end{aligned}$$

Hence

$$(9) \qquad \begin{aligned} (P_0^2 - D)\, z_{\tau-1} - Q_0 (P_0 - P_\tau)\, y_{\tau-1} &= Q_0 Q_\tau y_{\tau-2}, \\ Q_0 (P_0 + P_\tau)\, z_{\tau-1} - Q_0^2 y_{\tau-1} &= Q_0 Q_\tau z_{\tau-2}. \end{aligned}$$

According to the well-known relation $y_{\tau-1} z_{\tau-2} - y_{\tau-2} z_{\tau-1} = 1$†, we have just the formula

$$(10) \qquad (Q_0 y_{\tau-1} - P_0 z_{\tau-1})^2 - D z_{\tau-1}^2 = Q_0 Q_\tau.$$

Calling the period χ we have, from $Q_\tau = Q_\chi$, $Q_0 = Q_\chi$ and $P_0 = P_\chi$,

$$(11) \qquad (Q_\chi y_{\chi-1} - P_\chi z_{\chi-1})^2 - D z_{\chi-1}^2 = Q_\chi^2.$$

Further we obtain from (8″) $2 P_\chi z_{\chi-1} \equiv 0 \bmod. Q_\chi$ and, since D is without square factors, we get from (11) a solution of

$$x^2 - Dy^2 = (2^\varepsilon)^2. \qquad (\varepsilon = 0 \text{ or } 1)$$

where, in the case when $D \equiv 5 \bmod. 8$ and Q_χ is even, ε may be equal to unity. Finally we remark that for $\iota = \nu$ and on account of (5) the elements of the first category also are charactrized by

$$(12) \qquad D \equiv 0 \bmod. \frac{Q_\nu}{2^\varepsilon}. \qquad (\varepsilon = 0 \text{ or } 1)$$

* Perron, O.: Die Lehre von den Kettenbrüchen, p. 7.
† Perron, O., loc. cit., p. 16.

§ 2.

Referring to the formula (7) we are now able to construct systematically all periodically closed, symmetrical chains of fractions for a given D. We have therefore to determine P_0 and Q_0 according to

$$D = \mu\nu,\ Q_0 = 2^{\varepsilon}\mu,\ P_0 = \alpha\mu. \qquad (\varepsilon = 0 \text{ or } 1) \tag{13}$$

and to the inequalities $\alpha\mu > \sqrt{D}$, $(\alpha\mu - \sqrt{D})/\mu < 1$, hence α so that $\alpha\mu > \sqrt{D} > (\alpha - 1)\mu$ or

$$\alpha = \left[\frac{\sqrt{D}}{\mu}\right] + 1. \tag{14}$$

We must then prove $(\alpha\mu + \sqrt{D})/\mu$ to be a reduced fraction. But this is evident from $(\alpha\mu + \sqrt{D})/\mu > 1$. The possibility $\varepsilon = 1$ must always be discussed independently, giving $(\alpha\mu)^2 - \mu\nu = 2\mu\,[(\alpha^2\mu - \nu)/2]$. $0 < (\alpha\mu - \sqrt{D})/2\mu < 1$, $(\alpha\mu + \sqrt{D})/2\mu > 1$.

Since all chains with the initial element (13) have been constructed, the possible chain with $D = P_0^2 - Q_0^2$ and, as $\chi = 2\nu$, with $D = P_\nu^2 - Q_\nu^2$ are to be sought, which will completely determine all possible symmetric chains of reduced fractions. D being without square factors, $D = \mu\nu\ (\mu > \nu)$ implies $P_0 = (\mu + \nu)/2$, $Q_0 = (\mu - \nu)/2$ with μ and ν odd, and we have also

$$0 < \frac{\dfrac{\mu+\nu}{2} - \sqrt{D}}{\dfrac{\mu-\nu}{2}} < 1, \qquad \frac{\dfrac{\mu+\nu}{2} + \sqrt{D}}{\dfrac{\mu-\nu}{2}} > 1,$$

since $\mu + \nu - 2\sqrt{\mu\nu} < \mu - \nu$, $\mu > \nu$. Thus $(P_0 + \sqrt{D})/Q_0$ is a reduced fraction generating a periodically closed chain. Returning to (10) and (13) we may conclude from $(Q_0 y_{\rho-1} - P_0 z_{\rho-1})^2 - D z_{\rho-1}^2 = Q_0 Q_\rho$ that these symmetrical chains yield solutions of any equation $\mu x^2 - \nu y^2 = 2^{\varepsilon} Q_\rho$, because of the relation

$$\mu\,(2^{\varepsilon} y_{\rho-1} - \alpha z_{\rho-1})^2 - \nu z_{\rho-1}^2 = 2^{\varepsilon} Q_\rho, \tag{15}$$

which is valid for all denominators Q_ρ of the reduced fractions.

We shall now endeavour to determine the system of integers which form the periodic chain in the particular case $D = \mu\nu$, $P_0 = Q_0 = \mu > \nu$ or $D = \mu\nu$, $P_\nu = Q_\nu = \mu$. In this case the following reduced chain subsists

$$
\begin{aligned}
&\mu^2 - \mu r = \mu(\mu - r),\\
&(\mu - 2r)^2 - \mu r = (\mu - r)(\mu - 2^2 r),\\
(16')\qquad &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&[\mu - n(n-1)r]^2 - \mu r = [\mu - (n-1)^2 r][\mu - n^2 r],
\end{aligned}
$$

as far as $\mu - n^2 r > 0$, as may be checked by

$$
0 < \frac{\mu - n(n-1)r - \sqrt{D}}{\mu - n^2 r} < 1, \qquad \frac{\mu - n(n-1)r + \sqrt{D}}{\mu - n^2 r} > 1,
$$

remembering that $n\sqrt{r} < \sqrt{\mu}$. From the identity

$$
[\mu - n(n-1)r] + [\mu - n(n+1)r] = 2[\mu - n^2 r]
$$

we infer that all β_τ equal 2 as far as this formation holds. Also in the general case $\alpha \neq 1$ we can construct such sets

$$
\begin{aligned}
&(\alpha\mu)^2 - \mu r = \mu(\alpha^2\mu - r),\\
&[\alpha(2\alpha - 1)\mu - 2r]^2 - \mu r = (\alpha^2\mu - 1^2 r)((2\alpha - 1)^2\mu - 2^2 r),\\
(16'')\qquad &\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot\\
&\cdot\quad\cdot\quad\cdot\quad\cdot\quad\cdot
\end{aligned}
$$

$$
\begin{aligned}
&\{[\overline{n-1}\cdot\alpha - \overline{n-2}][n\alpha - \overline{n-1}]\mu - n(n-1)r\}^2 - \mu r\\
&\qquad = \{[\overline{n-1}\cdot\alpha - \overline{n-2}]^2\mu - (n-1)^2 r\}\{[n\alpha - \overline{n-1}]^2\mu - n^2 r\},
\end{aligned}
$$

where also, from the identity

$$
\begin{aligned}
&\{[\overline{n-1}\cdot\alpha - \overline{n-2}][n\alpha - \overline{n-1}]\mu - n(n-1)r\} + \{[n\alpha - \overline{n-1}]\\
&\qquad \cdot[\overline{n+1}\cdot\alpha - n]\mu - n(n+1)r\} = 2\{[n\alpha - \overline{n-1}]^2\mu - n^2 r\},
\end{aligned}
$$

all β_τ are equal to 2. It is readily seen that, when $\alpha = 1$, we recover from this (16′). Here also, when $[n\alpha - \overline{n-1}]\sqrt{\mu} - n\sqrt{r} > 0$, the inequality may be tested. For we have

$$
\begin{aligned}
&[\overline{n-1}\cdot\alpha - \overline{n-2}][n\alpha - \overline{n-1}]\mu - n(n-1)r - \sqrt{\mu r}\\
&= \{[n\alpha - \overline{n-1}]\sqrt{\mu} - n\sqrt{r}\}\{[\overline{n-1}\cdot\alpha - \overline{n-2}]\sqrt{\mu} + \overline{n-1}\cdot\sqrt{r}\} > 0,
\end{aligned}
$$

and from

$$
\{[n\alpha - \overline{n-1}]\sqrt{\mu} - n\sqrt{r}\}\{(\alpha - 1)\sqrt{\mu} + \sqrt{r}\} < 0,
$$

we have also

$$\{[n\alpha-\overline{n-1}]\sqrt{\mu}-n\sqrt{\nu}\}\{[\overline{n-1}\cdot\alpha-\overline{n-2}]\sqrt{\mu}+\overline{n-1}\cdot\sqrt{\nu}\}$$
$$<[n\alpha-\overline{n-1}]^2\mu-n^2\nu.$$

But again, since

$$\{[n\alpha-\overline{n-1}]\sqrt{\mu}+n\sqrt{\nu}\}\{[\overline{n-1}\cdot\alpha-\overline{n-2}]\sqrt{\mu}-\overline{n-1}\sqrt{\nu}\}$$
$$>[n\alpha-\overline{n-1}]^2\mu-n^2\nu,$$

it follows that

$$[n\alpha-\overline{n-1}][\overline{n-1}\cdot\alpha-\overline{n-2}]\mu-n(n-1)\nu+\sqrt{\mu\nu}$$
$$>[n\alpha-\overline{n-1}]^2\mu-n^2\nu,$$

and this, when $\sqrt{\nu}>(\alpha-1)\sqrt{\mu}$, viz. $\sqrt{D}>(\alpha-1)\mu$ (which is always satisfied), is equivalent to

$$[\overline{n-1}\cdot\alpha-\overline{n-2}]\sqrt{\mu}-\overline{n-1}\cdot\sqrt{\nu}>[n\alpha-\overline{n-1}]\sqrt{\mu}-n\sqrt{\nu}.$$

For $D=\mu\nu=P_i^2-Q_i^2$ ($i=0$, ν or $\nu+1$), we have $P_i=(\mu+\nu)/2$, $Q_i=(\mu-\nu)/2$, and the reduced system

(17)
$$\begin{aligned}
\left(\frac{\mu+\nu}{2}\right)^2-\mu\nu &= \frac{\mu-\nu}{2}\,\frac{\mu-\nu}{2},\\
\left(\frac{\mu-3\nu}{2}\right)^2-\mu\nu &= \frac{\mu-1^2\nu}{2}\,\frac{\mu-3^2\nu}{2},\\
&\cdots\cdots\\
&\cdots\cdots\\
\left(\frac{\mu-(2n-1)(2n+1)\nu}{2}\right)^2-\mu\nu &= \frac{\mu-(2n-1)^2\nu}{2}\,\frac{\mu-(2n+1)^2\nu}{2}.
\end{aligned}$$

We find in fact that

$$\frac{(\sqrt{\mu}-2n+1\sqrt{\nu})(\sqrt{\mu}+2n-1\sqrt{\nu})}{(\sqrt{\mu}-2n+1\sqrt{\nu})(\sqrt{\mu}+2n+1\sqrt{\nu})}<1,$$

supposing $\sqrt{\mu}>2n+1\cdot\sqrt{\nu}$, and

$$\frac{(\sqrt{\mu}+2n+1\sqrt{\nu})(\sqrt{\mu}-2n-1\sqrt{\nu})}{(\sqrt{\mu}+2n+1\sqrt{\nu})(\sqrt{\mu}-2n+1\sqrt{\nu})}>1.$$

Further, we find from

$$\frac{\mu-(2n-1)(2n+1)\nu}{2}+\frac{\mu-(2n+1)(2n+3)\nu}{2}=2\,\frac{\mu-(2n+1)^2\nu}{2},$$

that all β_τ are equal to 2 as far as this formation holds.

To illustrate the preceding we take $D = 105 = 3.5.7$ and consider first $Q_0 = 1$, whence, since $10 < \sqrt{105} < 11$, $P_0 = 11$, and then we have the following symmetrical chain of fractions

$$
1)\quad
\begin{aligned}
\frac{11+\sqrt{105}}{1} &= 22 - \frac{11-\sqrt{105}}{1} = 22 - \cfrac{1}{\cfrac{11+\sqrt{105}}{16}}, \\
\frac{21+\sqrt{105}}{21} &= 2 - \frac{21-\sqrt{105}}{21} = 2 - \cfrac{1}{\cfrac{21+\sqrt{105}}{16}}, \\
\frac{11+\sqrt{105}}{16} &= 2 - \frac{21-\sqrt{105}}{16} = 2 - \cfrac{1}{\cfrac{21+\sqrt{105}}{21}}, \\
\frac{21+\sqrt{105}}{16} &= 2 - \frac{11-\sqrt{105}}{16} = 2 - \cfrac{1}{\cfrac{11+\sqrt{105}}{1}}.
\end{aligned}
$$

The period $\chi = 2\nu = 4$ and thus $P_2 = Q_2 = 21$ giving $D = 105 \equiv 0$ mod. 21. It may be remarked that the chain 1) is formed by the values of the system (16″). Further the continued fraction is, as it should be, symmetrical on both sides of $\beta_2 = 2$, $\beta_0 = 22$ excepted. This type of chain ($Q_0 = 1$), the principal chain, is commonly used to solve Pell's equation.

2) The case $\varepsilon = 1$ will give $Q_0 = 2$, $P_0 = 11$.

$$
\begin{aligned}
\frac{11+\sqrt{105}}{2} &= 11 - \frac{11-\sqrt{105}}{2} = 11 - \cfrac{1}{\cfrac{11+\sqrt{105}}{8}}, \\
\frac{13+\sqrt{105}}{8} &= 3 - \frac{11-\sqrt{105}}{8} = 3 - \cfrac{1}{\cfrac{11+\sqrt{105}}{2}}, \\
\frac{11+\sqrt{105}}{8} &= 3 - \frac{13-\sqrt{105}}{8} = 3 - \cfrac{1}{\cfrac{13+\sqrt{105}}{8}}.
\end{aligned}
$$

The period $\chi = 2\nu + 1 = 3$, $105 = P_2^2 - Q_2^2 = 13^2 - 8^2$ from $D = \mu\nu = 21.5$.

3) $Q_0 = 3$, $P_0 = 4.3$, found by $3\alpha > \sqrt{105} > 3(\alpha - 1)$.

$$\frac{12+\sqrt{105}}{3} = 8 - \frac{12-\sqrt{105}}{3} = 8 - \cfrac{1}{\cfrac{12+\sqrt{105}}{13}},$$

$$\frac{14+\sqrt{105}}{7} = 4 - \frac{14-\sqrt{105}}{7} = 4 - \cfrac{1}{\cfrac{14+\sqrt{105}}{13}},$$

$$\frac{12+\sqrt{105}}{13} = 2 - \frac{14-\sqrt{105}}{13} = 2 - \cfrac{1}{\cfrac{14+\sqrt{105}}{7}},$$

$$\frac{14+\sqrt{105}}{13} = 2 - \frac{12-\sqrt{105}}{13} = 2 - \cfrac{1}{\cfrac{12+\sqrt{105}}{3}}.$$

The period $2\nu = 4$, $Q_2 = 7$, $P_2 = 14$, $D = 105 \equiv 0$ mod. 7, and the chain is formed partly like (16″).

4) $Q_0 = 2.3$, $P_0 = 5.3$ by $3\alpha > \sqrt{105} > 3(\alpha - 2)$.

$$\frac{15+\sqrt{105}}{6} = 5 - \frac{15-\sqrt{105}}{6} = 5 - \cfrac{1}{\cfrac{15+\sqrt{105}}{20}},$$

$$\frac{27+\sqrt{105}}{24} = 2 - \frac{21-\sqrt{105}}{24} = 2 - \cfrac{1}{\cfrac{21+\sqrt{105}}{14}},$$

$$\frac{15+\sqrt{105}}{20} = 2 - \frac{25-\sqrt{105}}{20} = 2 - \cfrac{1}{\cfrac{25+\sqrt{105}}{26}},$$

$$\frac{21+\sqrt{105}}{14} = 3 - \frac{21-\sqrt{105}}{14} = 3 - \cfrac{1}{\cfrac{21+\sqrt{105}}{24}},$$

$$\frac{25+\sqrt{105}}{26} = 2 - \frac{27-\sqrt{105}}{26} = 2 - \cfrac{1}{\cfrac{27+\sqrt{105}}{24}}, \text{ \&c.}$$

The period $\chi = 2\nu = 8$, $Q_4 = 2.7$, $P_4 = 3.7$, $D = 105 \equiv 0$ mod. $Q_4/2$.
5) $Q_0 = 5$, $P_0 = 3.55\alpha > \sqrt{105} > 5(\alpha - 1)$.

We find the period $\chi = 2\nu = 20$, $P_{10} = Q_{10} = 105$, $D = 105 \equiv 0$ mod. 105. We have here a very interesting example of the formation (16″) and at the

The first part of the theorem is obvious. For each chain (2) must run into a periodically closed chain; and by Theorem 7 we get a solution of (1) when, and only when, the chain is the principal chain, viz., when it contains the denominator $Q_0 = 1$. The second part we can prove as follows. We start from the pair of given relations

$$M_1^2 - DN_1^2 = A, \qquad M_2^2 - DN_2^2 = A,$$

where M_1, N_1 and M_2, N_2 are relatively prime, and both solutions belong to the same congruence-root $\alpha_1^{(1)}$. With $\alpha_0^{(1)} + \alpha_1^{(1)} = A$ we form

$$\frac{\alpha_0^{(1)} + \sqrt{D}}{A} = 1 - \frac{\alpha_1^{(1)} - \sqrt{D}}{A} = 1 - \frac{1}{\dfrac{\alpha_1^{(1)} + \sqrt{D}}{A_1}}, \text{ \&c.}$$

From (10) § 1 we get

$$(Ay_\rho - \alpha_0^{(1)} z_\rho)^2 - Dz_\rho^2 = 1 \cdot A,$$

and therefore, with $N_1 = z_\rho$, $M_1 = \pm(Ay_\rho - \alpha_0^{(1)} z_\rho)$, the congruence

(3) $$M_1 \equiv \pm \alpha_1^{(1)} N_1 \text{ mod. } A, \text{ and also } M_2 \equiv \pm \alpha_1^{(1)} N_2 \text{ mod. } A.$$

Forming the ratio

(4) $$\begin{aligned} \frac{M_1 + N_1\sqrt{D}}{M_2 \pm N_2\sqrt{D}} &= \frac{(M_1 + N_1\sqrt{D})(M_2 \mp N_2\sqrt{D})}{A} \\ &= \frac{M_1M_2 \mp DN_1N_2}{A} + \sqrt{D}\,\frac{M_2N_1 \mp M_1N_2}{A} = \beta + \gamma\sqrt{D}, \end{aligned}$$

we find by (3)

(5) $$M_1M_2 \equiv \pm \alpha_1^{(1)2} N_1N_2, \qquad M_1M_2 \equiv \pm DN_1N_2 \text{ mod. } A,$$

and also

$$M_2N_1 \equiv \pm M_1N_2 \text{ mod. } A,$$

whence $\beta + \gamma\sqrt{D}$ must be a unit, and Theorem 8 is proved. As a corollary we have the solution of

(6) $$ax^2 + 2bxy + cy^2 = A$$

in integers x and y, Gauss's well-known theorem. We form

(7) $$(ax + by)^2 - Dy^2 = aA$$

with any solution

(8) $$M_\nu^2 - DN_\nu^2 = aA$$

belonging to some congruence-root $z = \alpha_1^{(\nu)}$, $\alpha_1^{(\tau)}$ $(\tau = 1 \cdots \nu \cdots)$ being the roots of $z^2 \equiv D \bmod. aA$. Thus we have $\alpha_1^{(\tau)2} \equiv D$, $b^2 \equiv D \bmod. a$ and at least for some one of $\alpha_1^{(\tau)}$ the congruence must hold, say $\alpha_1^{(\nu)}, \alpha_1^{(\nu)} \equiv b \bmod. a$. By (3) we get then either $M_\nu + bN_\nu \equiv 0$ or $M_\nu - bN_\nu \equiv 0 \bmod. a$. With $y = \pm N_\nu$ we must therefore always obtain x in $ax \pm bN_\nu = M_\nu$ as an integer, whence we must seek all solutions of (6) by forming the chains (2) starting from all congruence-roots $\alpha_1^{(i)}$ $(i = 1, 2 \cdots \tau)$ mod. aA such that

$$\alpha_1^{(p)} \equiv \pm b \bmod. a.$$

If A is a negative integer the chain terminates

$$\alpha_{i-1}^{(1)2} - D = A_{i-2}A_{i-1}, \qquad \alpha_i^{(1)2} - D = A_{i-1}A_i,$$

where $A_{i-1} = D - 1$ and $A_i = D$ by the properties of the closed chains of positively reduced fractions, given in the preceding paragraphs. As can be seen from (2), all A_τ in the chain may also be expressed in the principal form (1), and the integers x and y may be calculated by the continued fraction formed. But, conversely, since the solution

$$(9') \qquad M_1^2 - DN_1^2 = A$$

is given,

$$(9'') \qquad M_2^2 - DN_2^2 = A_1$$

may easily be computed. For, since $\alpha_1^{(1)2} - D = AA_1$ and since $M_1 \equiv \pm N_1 \alpha_1^{(1)}$, $M_1\alpha_1^{(1)} \equiv \pm N_1\alpha_1^{(1)2} \equiv \pm N_1 D$ mod. A, the ratio of

$$\frac{(\alpha_1^{(1)} + \sqrt{D})(M_1 \mp N_1\sqrt{D})}{A} = \frac{M_1\alpha_1^{(1)} \mp N_1 D}{A} + \sqrt{D}\,\frac{M_1 \mp N_1\alpha_1^{(1)}}{A}$$

will at once give us M_2 and N_2 as integers in (9'').

Example.

$$6x^2 - 20xy + 11y^2 = 99, \quad (6x - 10y)^2 - 34y^2 = 6.99.$$

We have mod. 594 the four roots ± 230, ± 122 and have $230 \equiv -10 \bmod. 6$ and also $122 \equiv -10 \bmod. 6$. Hence

$$230^2 - 34 = 594.89, \quad 37^2 - 34 = 89.15, \quad 8^2 - 34 = 15.2, \quad 6^2 - 34 = 2.1,$$

and, from

$$\frac{6+\sqrt{34}}{1} = 12 - \frac{6-\sqrt{34}}{1} = 12 - \frac{1}{\frac{6+\sqrt{34}}{2}},$$

$$\frac{8+\sqrt{34}}{15} = 3 - \frac{37-\sqrt{34}}{15} = 3 - \frac{1}{\frac{37+\sqrt{34}}{89}},$$

$$\frac{6+\sqrt{34}}{2} = 7 - \frac{8-\sqrt{34}}{2} = 7 - \frac{1}{\frac{8+\sqrt{34}}{15}},$$

$$\frac{37+\sqrt{34}}{89} = 3 - \frac{230-\sqrt{34}}{89} = 3 - \frac{1}{\frac{230+\sqrt{34}}{594}},$$

we calculate

$$12 - \frac{1|}{|2} - \frac{1|}{|3} - \frac{1|}{|3} = \frac{628}{53},$$

having

$$(628 - 6.53)^2 - 34.53^2 = 594.$$

Taking $y = 53$ we find $x = 140$ from equation $6x - 10.53 = 310$, but in view of $35^2 - 34.6^2 = 1$ we deduce the simpler solution $x = 2$, $y = 5$. Proceeding from the root 122 we form the chain

$$122^2 - 34 = 594.25,\ 28^2 - 34 = 25.30,\ 32^2 - 34 = 30.33,$$
$$34^2 - 34 = 33.34,$$

leading to

$$\frac{34+\sqrt{34}}{34} = 2 - \frac{34-\sqrt{34}}{34} = 2 - \frac{1}{\frac{34+\sqrt{34}}{33}},$$

$$\frac{32+\sqrt{34}}{30} = 2 - \frac{28-\sqrt{34}}{30} = 2 - \frac{1}{\frac{28+\sqrt{34}}{25}},$$

$$\frac{34+\sqrt{34}}{33} = 2 - \frac{32-\sqrt{34}}{33} = 2 - \frac{1}{\frac{32+\sqrt{34}}{30}},$$

$$\frac{28+\sqrt{34}}{25} = 6 - \frac{122+\sqrt{34}}{25} = 6 - \frac{1}{\frac{122+\sqrt{34}}{594}},$$

and have

$$2 - \frac{1|}{|2} - \frac{1|}{|2} - \frac{1|}{|6} = \frac{21}{16},$$

giving the solution

$$34\,(21 - 16)^2 - 16^2 = 594.$$

Since $y = 5$ and $6x - 10.5 = 16$, we have $x = 11$ as a solution of

$$6x^2 - 20xy + 11y^2 = -99$$

giving a check.

§ 4.

This paragraph will be devoted to the continued fractions, which appear in the principal chain $Q_0 = 1$. The main results may be summarized in the following theorem.

THEOREM 9. *In the principal chain having the period $\chi = 2\nu$ the middle element is as usual characterized by*

(1) $$D \equiv 0 \quad \text{mod. } Q_\nu/2^\varepsilon \qquad (\varepsilon = 0, 1)$$

and the fraction $y_{\nu-1}/z_{\nu-1}$ leads to a solution α_1, β_1 of the equation

(2) $$\mu x^2 - \nu y^2 = 2^\varepsilon$$

with $D = \mu\nu$. Then the solution α_2, β_2 of Pell's equation is given by means of equation

(3) $$\alpha_2 + \beta_2\sqrt{D} = \frac{1}{2^\varepsilon}\,(\alpha_1\sqrt{\mu} + \beta_1\sqrt{\nu})^2.$$

In the case $\chi = 2\nu + 1$ the middle element must be of the second category and it is always characterized by

(4) $$D = [(D+1)/2]^2 - [(D-1)/2]^2.$$

Being given α_2 and β_2, we find first, in virtue of

(5) $$\begin{aligned} \sqrt{\alpha_2 + \beta_2\sqrt{D}} + \sqrt{\alpha_2 - \beta_2\sqrt{D}} &= \sqrt{2(\alpha_2+1)}, \\ \sqrt{\alpha_2 + \beta_2\sqrt{D}} - \sqrt{\alpha_2 - \beta_2\sqrt{D}} &= \sqrt{2(\alpha_2-1)}, \end{aligned}$$

that

(6) $$2\sqrt{\alpha_2 + \beta_2\sqrt{D}} = \sqrt{2(\alpha_2+1)} + \sqrt{2(\alpha_2-1)}.$$

When α_2 is odd we must get $\sqrt{2(\alpha_2+1)} = 2\alpha_1\sqrt{\mu}$, $\sqrt{2(\alpha_2-1)} = 2\beta_1\sqrt{\nu}$ giving a solution of $\mu\alpha_1^2 - \nu\beta_1^2 = 1$, and when α_2 is even a solution of $\mu\alpha_1^2 - \nu\beta_1^2 = 2$. And now from (6) we immediately obtain (3). We shall treat as particularly interesting the case $\mu = D$, $\nu = 1$, α_1 and β_1 leading to a solution of $\alpha_1^2 - D\beta_1^2 = -1$. But the solution that we must take from the chain must be that of $(D\beta_1)^2 - D\alpha_1^2 = D$, whence $Q_\nu = D$, and about this element a formation (16′) groups itself for $\mu = D$, $\nu = 1$. With D all $D - n^2 > 0$ $(n = 1 \cdots r)$ must also be denominators of reduced fractions belonging to the chain, whence we have the corollary: *The equation $x^2 - Dy^2 = -1$ being soluble, D itself may be expressed in the form $\alpha^2 + \beta^2 - D\gamma^2$ and the number n of such essentially different expressions equals the largest n given by $D - n^2 > 0$.* Conversely, α_1 and β_1 being given so that $\alpha_1^2 - D\beta_1^2 = -1$ we compute

$$\frac{(D - \sqrt{D})(D\beta_1 + \alpha_1\sqrt{D})}{D} = (D\beta_1 - \alpha_1) + \sqrt{D}(\alpha_1 - \beta_1).$$

and $(D\beta_1 - \alpha_1)^2 - D(\alpha_1 - \beta_1)^2 = D - 1$, which is easily verified. Having also $(D\beta_1)^2 - D\alpha_1^2 = D$, and because of the well-known relations*

$$y_\tau = \beta_\tau y_{\tau-1} - y_{\tau-2}, \qquad z_\tau = \beta_\tau z_{\tau-1} - z_{\tau-2}, \tag{7}$$

and the formula (10) of § 1 for $Q_0 = 1$, $(1 y_{\tau-1} - P_0 z_{\tau-1})^2 - D z_{\tau-1}^2 = Q_\tau$ and $D\beta_1(\alpha_1 - \beta_1) - \alpha_1(D\beta_1 - \alpha_1) = -1$,† we can directly calculate x_2 and y_2, solving $x_2^2 - Dy_2^2 = D - 2^2$ and getting, since $\beta_\tau = 2$,

$$D\beta_1 = 2(D\beta_1 - \alpha_1) - x_2, \qquad \alpha_1 = 2(\alpha_1 - \beta_1) - y_2. \tag{8}$$

In the same way the following x_i and y_i belonging to the denominators $D - i^2$ may be successively computed, which is rather convenient as the continued fractions themselves have often a very long period χ. It may however be observed that the part of the continued fraction with $\beta_\tau = 2$ may easily be reduced, since

$$\underbrace{\frac{1|}{|2} - \frac{1|}{|2} - \cdots - \frac{1|}{|2}}_{\nu} = \frac{\nu}{\nu + 1}. \tag{9}$$

* Perron, O.: loc. cit. p. 5.

† This shows that these two solutions are formed from two successive fractions $y_{\tau-1}/z_{\tau-1}$, y_τ/z_τ in the same part of the reduced chain, which can be proved something like Theorem 13, Perron, loc. cit. p. 5.

by a simple induction. In the case of the odd period $\chi = 2\nu + 1$ we must have $D = P_{\nu+1}^2 - Q_{\nu+1}^2$. In view of (6), and as (2) must not be satisfied since that would lead to Q_ν as denominator of a fraction of the first category, the only possibility left will be

$$\alpha_1^2 - D\beta_1^2 = -2. \tag{10}$$

As $Q_\nu = -2$, a negative number, does not appear in the chain, we have to look for the Q_ν actually found in the chain. From (6) we conclude α_2 to be even in order to give (10), and thus D is odd and hence also α_1, β_1. Thus $1 - D = -2$, $D = 3$ mod. 4, whence $(D-1)/2$ is odd. Further the ratio of $1^2 - 1^2 D = -(D-1)$ and, since

$$\frac{(\alpha_1 + \beta_1 \sqrt{D})(1 + 1\sqrt{D})}{2} = \frac{D\beta_1 + \alpha_1}{2} + \sqrt{D}\,\frac{\alpha_1 + \beta_1}{2}, \tag{11}$$

$\alpha_1^2 - D\beta_1^2 = -2$ will leave the equation

$$\left(\frac{D\beta_1 + \alpha_1}{2}\right)^2 - D\left(\frac{\alpha_1 + \beta_1}{2}\right)^2 = \frac{D-1}{2}. \tag{12}$$

Supposing $(D-1)/2$ and $(\alpha_1 + \beta_1)$ to have any common factor m, we should find from $D = 1$ mod. $(D-1)/2\alpha_1^2 - \beta_1^2 = -2$, $2 = 0$ mod. m, which for m odd is impossible. On account of the congruence

$$\frac{D\beta_1 + \alpha_1}{2} = \frac{D+1}{2}\,\frac{\alpha_1 + \beta_1}{2} \qquad \text{mod.}\ \frac{D-1}{2},$$

we get, by (12), $[(D+1)/2]^2 = D$ mod. $(D-1)/2$, whence this solution (12) must belong to the congruence-root $\pm(D+1)/2$. Thus the reduced fraction $\dfrac{\frac{D+1}{2} + \sqrt{D}}{\frac{D-1}{2}}$ must appear in the principal chain and its middle element must therefore be formed from $D = [(D+1)/2]^2 - [(D-1)/2]^2$, as is required by the statement in Theorem 9. Further the formation (17) for $\mu = D$, $\nu = 1$ will reach as far as $D - (2n+1)^2 > 0$. Conversely, having by the continued fraction a solution $\alpha_\nu^2 - D\beta_\nu^2 = (D-1)/2$, it will be a simple matter to find a solution of $x^2 - Dy^2 = -2$. As α_ν, β_ν must belong to the root $\pm(D+1)/2$, we get

$$\alpha_\nu = \pm\frac{D+1}{2}\beta_\nu \qquad \text{mod.}\ \frac{D-1}{2};$$

hence either $\alpha_\nu + \beta_\nu \equiv 0$ or $\alpha_\nu - \beta_\nu \equiv 0 \bmod.(D-1)/2$. Let us assume the latter case and therefore also $\alpha_\nu - D\beta_\nu \equiv 0 \bmod.(D-1)/2$. We may then form the ratio

$$(13)\qquad \frac{(1+\sqrt{D})(\alpha_\nu - \sqrt{D}\beta_\nu)}{\frac{D-1}{2}} = \frac{\alpha_\nu - D\beta_\nu}{\frac{D-1}{2}} + \sqrt{D}\,\frac{\alpha_\nu - \beta_\nu}{\frac{D-1}{2}},$$

and we have in this way gained a solution of $x^2 - Dy^2 = -2$.

Example. $D = 51$.

$$\frac{8+\sqrt{51}}{1} = 16 - \frac{8-\sqrt{51}}{1} = 16 - \frac{1}{\frac{8+\sqrt{51}}{13}},$$

$$\frac{18+\sqrt{51}}{21} = 2 - \frac{24-\sqrt{51}}{21} = 2 - \frac{1}{\frac{24+\sqrt{51}}{25}},$$

$$\frac{8+\sqrt{51}}{13} = 2 - \frac{18-\sqrt{51}}{13} = 2 - \frac{1}{\frac{18+\sqrt{51}}{21}},$$

$$\frac{24+\sqrt{51}}{25} = 2 - \frac{26-\sqrt{51}}{25} = 2 - \frac{1}{\frac{26+\sqrt{51}}{25}} = \&c.$$

Period $2\nu + 1 = 7$. We verify in fact $(61 - 8.4)^2 - 51.4^2 = 25$, and by (13) $4.51 - 29 = 7.25$, $4 - 29 = -1.25$ leading to the solution of $7^2 - 51.1^2 = -2$.

§ 5.

In this paragraph we must demonstrate the connection between the chains of the positively and negatively reduced fractions, in order to show their relations to Gauss's classical theory of quadratic forms. We shall first prove this theorem.

THEOREM 1. *From a periodically closed chain of positive type it is always possible to pass over directly to a chain of negative type.*

In all periodic chains, as we have seen, there exist Q_τ with $-P_\tau \equiv \alpha_\tau$ mod. Q_τ and $\alpha_\tau^2 - D < 0$. Putting $(P_\tau + \sqrt{D})/Q_\tau = \mu_\tau + (\sqrt{D} - \alpha_\tau)/Q_\tau$, we see that, either μ_τ or possibly $\nu_\tau > \mu_\tau$ being the greatest integer in $(P_\tau + \sqrt{D})/Q_\tau$, there results

$$\frac{P_\tau + \sqrt{D}}{Q_\tau} = \nu_\tau + \frac{\sqrt{D} - [\alpha_\tau + (\nu_\tau - \mu_\tau)Q_\tau]}{Q_\tau},$$

$$\sqrt{D} - [\alpha_\tau + (\nu_\tau - \mu_\tau)Q_\tau] > 0.$$

since the fraction must be added. In all cases there exists an equality

$$\frac{P_\tau+\sqrt{D}}{Q_\tau}=r_\tau+\frac{\sqrt{D}-P_{\tau+1}}{Q_\tau}=r_\tau+\cfrac{1}{\cfrac{\sqrt{D}+P_{\tau+1}}{Q_{\tau+1}}},$$

with $P_{\tau+1}>0$ and $0<(\sqrt{D}-P_{\tau+1})/Q_\tau<1$. We can also assume the element $(P_\tau+\sqrt{D})/Q_\tau$ to be chosen so that in $P_\tau^2-D=Q_{\tau-1}Q_\tau$ we still have $Q_{\tau-1}>Q_\tau$. Combining this with $D-P_{\tau+1}^2=Q_\tau Q_{\tau+1}$ we have

$$P_\tau^2-P_{\tau+1}^2=(P_\tau-P_{\tau+1})(P_\tau+P_{\tau+1})=Q_\tau(Q_{\tau-1}+Q_{\tau+1})>Q_\tau^2.$$

If $P_\tau+P_{\tau+1}=Q_\tau$, then $P_\tau-P_{\tau+1}>Q_\tau$, which is impossible. Therefore we must have $P_\tau+P_{\tau+1}>2Q_\tau$ and $r_\tau\geq 2$ and thus $(\sqrt{D}+P_{\tau+1})/Q_\tau=r_\tau-(P_\tau-\sqrt{D})/Q_\tau>1$, whence $(\sqrt{D}+P_{\tau+1})/Q_\tau$ must be a negatively reduced fraction. Concluding further, in the same way as for positively reduced fractions, that they must all exist within the periodic chains, our Theorem 1 is proved. As a corollary we get: *In all periodic chains of positive type there exist elements $Q_\tau<2\sqrt{D}$*, as can be seen from $0<(\sqrt{D}-P_{\tau+1})/Q_\tau<1$, $(\sqrt{D}+P_{\tau+1})/Q_\tau>1$, giving $Q_\tau<\sqrt{D}+P_{\tau+1}<2\sqrt{D}$.

THEOREM 2. *The complementary chain of a positively reduced fraction is always connected with the complementary chain of the negatively reduced fraction.*

We have $(P_\rho+\sqrt{D})/Q_\rho=\beta_\rho-(P_{\rho+1}-\sqrt{D})/Q_\rho$, $(P_\rho+\sqrt{D})/Q_\rho=r_\rho+(\sqrt{D}-r_\rho)/Q_\rho$, $(\sqrt{D}+r_\rho)/Q_\rho=\delta_\rho+(\sqrt{D}-r_{\rho+1})/Q_\rho$ and thus $P_{\rho+1}+P_\rho\equiv 0$, $P_\rho+r_\rho\equiv 0$, $r_{\rho+1}+r_\rho\equiv 0 \bmod. Q_\rho$, whence $P_{\rho+1}+r_{\rho+1}\equiv 0 \bmod. Q_\rho$, and the theorem is proved.

Two fractions ω_ν and ω_τ between which the relation

$$\omega_\nu=\frac{\alpha\,\omega_\tau+\beta}{\gamma\,\omega_\tau+\delta}$$

exists are said, to be properly equivalent when $\alpha\delta-\beta\gamma=+1$ and improperly equivalent when $\alpha\delta-\beta\gamma=-1$.

THEOREM 3. *The same definition of the negatively reduced fractions is valid also for Gauss's reduced forms. Thus each form gives rise to a reduced fraction, and each form can also be represented by a single fraction.*

By $0<(\sqrt{D}-P_\rho)/Q_\rho<1$, $(\sqrt{D}+P_\rho)/Q_\rho>1$ it follows that $\sqrt{D}-P_\rho<Q_\rho<\sqrt{D}+P_\rho$, $0<P_\rho<\sqrt{D}$, and in this same way Gauss has defined the two reduced forms $(Q_\rho, P_\rho, -Q_{\rho-1})$ and $(-Q_\rho, P_\rho, Q_{\rho-1})$.* Now let $(-Q_\rho, P_\rho,$

* Gauss, C. F.: Untersuchungen über Höhere Mathematik (Deutsch v. Maser, H.), 1889, p. 152.

$Q_{\rho-1}$) be represented by the reduced fraction $(\sqrt{D}+P_\rho)/-Q_\rho$. Then we have realized a one-to-one correspondence between forms and fractions. It may be observed that proceeding from these fractions we can always form chains leading only to substitutions of proper equivalence (determinant $+1$), as follows

$$\frac{9+\sqrt{105}}{-6} = (-3) - \frac{9-\sqrt{105}}{-6} = (-3) - \cfrac{1}{\cfrac{9+\sqrt{105}}{4}},$$

$$\frac{7+\sqrt{105}}{-14} = (-1) - \frac{7-\sqrt{105}}{-14} = (-1) - \cfrac{1}{\cfrac{7+\sqrt{105}}{4}},$$

$$\frac{9+\sqrt{105}}{4} = 4 - \frac{7-\sqrt{105}}{4} = 4 - \cfrac{1}{\cfrac{7+\sqrt{105}}{-14}},$$

$$\frac{7+\sqrt{105}}{4} = 4 - \frac{9-\sqrt{105}}{4} = 4 - \cfrac{1}{\cfrac{9+\sqrt{105}}{-6}}.$$

Following up the parallel between Gauss's theory and this theory of periodic chains, we shall find that Gauss's method of forming adjacent forms* is exactly the same as our method of forming a chain, which the reader will have no difficulty in seeing from a numerical example, say $D = 79$.† Thus, since these reduced fractions with negative denominator can always be brought over into reduced chains of positive type, we have achieved the fundamental.

THEOREM 4. *The number of Gauss's classes equals the number of positively periodic chains.*

Further we shall prove this important theorem.

THEOREM 5. *If the fundamental unity* η *has a negative norm* $(N(\eta) = -1)$, *all periodically closed chains of negative type will have an odd period, but, if* $N(\eta) = +1$, *all periods will be even.*

Proceeding from the negatively reduced fraction $(\sqrt{D}+P_\rho)/Q_\rho$ and forming the chain as above, this chain will be closed with $(\sqrt{D}+P_\rho)/Q_\rho$ in all chains of even period but with $(\sqrt{D}+P_\rho)/-Q_\rho$ in the case of an odd period. In this latter case we take by Theorem 7 § 1 a solution of $x^2 - Dy^2 = -(2^\epsilon)^2$ ($\epsilon = 0$ or 1). Thus we have either $N(\eta) = -1$ or $\alpha^2 - D\beta^2 = -4$, $N(\eta) = +1$, which is impossible. For we find $\lambda^3 = [(\alpha+\beta\sqrt{D})/2]^3 = [(\alpha-\beta)/2 + \beta(\sqrt{D}+1)/2]^3 = m + n\sqrt{D}$ (m, n integers) from $D = 5$

* Gauss, C. F., loc. cit., p. 118.

† Gauss, C. F., loc. cit., p. 159.

mod. 8. Hence $N(\lambda^3) \doteq -1$ contrary to $N(\eta) = +1$. Thus, with only one of the periods odd there must result $N(\eta) = -1$. But then all forms $(-a, b, c)$ and $(a, b, -c)$ are properly equivalent* and hence also the two reduced fractions $(P_\rho + \sqrt{D})/Q_\rho$ and $(P_\rho + \sqrt{D})/-Q_\rho$, whence they stay in a reduced chain of odd period. Thus for $N(\eta) = -1$ all closed chains have necessarily an odd period and for $N(\eta) = +1$ they have an even period. Hence we get the corollary: *If* $N(\eta) = +1$, *the number of positively reduced chains is twice that of the negatively reduced chain, but if* $N(\eta) = -1$ *these numbers are equal.*

THEOREM 6. *Ambiguous classes of forms*† *correspond to symmetrically closed chains with elements* $(\sqrt{D} + P_\nu)/Q_\nu$ *of the first category*

$$D \equiv 0 \bmod. Q_\nu/2^\epsilon.$$

This follows immediately from Gauss's definition of ambiguous forms and from (5) § 1, it being observed that for $D > 0$ every ambiguous class of forms certainly contains two ambiguous forms.‡ But given the form $(Q_\rho, P_\rho, -Q_{\rho-1})$ with $Q_{\rho-1} = Q_\rho$, there exists a corresponding symmetrical chain without any element of the first category and therefore only of the second if $N(\eta) = +1$.

Example. $D = 34$, $34 = 3^2 + 5^2$.

$$\frac{\sqrt{34}+3}{5} = 1 + \frac{\sqrt{34}-2}{5} = 1 + \cfrac{1}{\cfrac{\sqrt{34}+2}{6}},$$

$$\frac{\sqrt{34}+5}{3} = 3 + \frac{\sqrt{34}-4}{3} = 3 + \cfrac{1}{\cfrac{\sqrt{34}+4}{6}},$$

$$\frac{\sqrt{34}+2}{6} = 1 + \frac{\sqrt{34}-4}{6} = 1 + \cfrac{1}{\cfrac{\sqrt{34}+4}{3}},$$

$$\frac{\sqrt{34}+4}{6} = 1 + \frac{\sqrt{34}-2}{6} = 1 + \cfrac{1}{\cfrac{\sqrt{34}+2}{5}},$$

$$\frac{\sqrt{34}+3}{3} = 3 + \frac{\sqrt{34}-5}{3} = 3 + \cfrac{1}{\cfrac{\sqrt{34}+5}{3}},$$

$$\frac{\sqrt{34}+2}{5} = 1 + \frac{\sqrt{34}-3}{5} = 1 + \cfrac{1}{\cfrac{\sqrt{34}+3}{5}}.$$

* Sommer, J.: Vorlesungen über Zahlentheorie, p. 203.

† Gauss, C. F., loc. cit., p. 125.

‡ Gauss, C. F., loc. cit., p. 215 and art. 194.

Then we remark that symmetric chains of positive type must always pass over into symmetric chains of negative type. But for $N(\eta) = +1$ there exists a symmetrical negative chain having two elements of the second category and to this chain there cannot correspond any symmetric positive chain, since the two derived forms are improperly but never properly equivalent.

Example. $D = 34$.

$$\frac{7+\sqrt{34}}{3} = 4 + \frac{\sqrt{34}-5}{3} = 4 + \cfrac{1}{\frac{\sqrt{34}+5}{3}},$$

$$\frac{12+\sqrt{34}}{11} = 2 - \frac{10-\sqrt{34}}{11} = 2 - \cfrac{1}{\frac{10+\sqrt{34}}{6}},$$

$$\frac{7+\sqrt{34}}{3} = 5 - \frac{8-\sqrt{34}}{3} = 5 - \cfrac{1}{\frac{8+\sqrt{34}}{10}},$$

$$\frac{10+\sqrt{34}}{6} = 3 - \frac{8-\sqrt{34}}{6} = 3 - \cfrac{1}{\frac{8+\sqrt{34}}{5}},$$

$$\frac{8+\sqrt{34}}{10} = 2 - \frac{12-\sqrt{34}}{10} = 5 - \cfrac{1}{\frac{12+\sqrt{34}}{11}},$$

$$\frac{8+\sqrt{34}}{5} = 3 - \frac{7-\sqrt{34}}{5} = 3 - \cfrac{1}{\frac{7+\sqrt{34}}{3}}.$$

By Theorem 4 it must be possible to establish a one-to-one correspondence between positively reduced fractions appertaining to $\sqrt{D}$ and quadratic forms $(Q_\rho, -P_\rho, Q_{\rho-1})$ of determinant D. This is accomplished by choosing as correspondents $(P_\rho + \sqrt{D})/Q_\rho$ and $(Q_\rho, -P_\rho, Q_{\rho-1})$. But the definition of a reduced form will now be distinct from that of Gauss, the form $(Q_\rho, -P_\rho, Q_{\rho-1})$ being reduced for values Q_ρ, P_ρ from $P_\rho - \sqrt{D} < Q_\rho < P_\rho + \sqrt{D}$, $\sqrt{D} < P_\rho < D$.

Thus for example we see from

$$\frac{-15+\sqrt{105}}{5} = 0 - \frac{15-\sqrt{105}}{5} = 0 - \cfrac{1}{\frac{15+\sqrt{105}}{24}},$$

that the substitution of determinant $+1$, $x = 0x' - 1y'$, $y = 1x' + 0y'$, transforms $5x^2 + 30xy + 24y^2$ into the reduced form $24x'^2 - 30x'y' + 5y'^2$, and from

$$\frac{15+\sqrt{105}}{-5}=(-4)-\frac{\sqrt{105}-5}{5}=(-4)-\cfrac{1}{\cfrac{\sqrt{105}+5}{16}},$$

$$\frac{5+\sqrt{105}}{16}=\quad 1-\frac{11-\sqrt{105}}{16}=\quad 1-\cfrac{1}{\cfrac{11+\sqrt{105}}{1}},$$

that the substitution of proper equivalence $x=-5x'+4y'$, $y=1x'-1y'$ transforms $-5x^2-30xy-24y^2$ into the reduced form $x'^2-22x'y'+16y'^2$. In this connection we may also point out the disadvantage that in this new definition of reduced forms ambiguous classes can exist without reduced ambiguous forms, as we see from 8) § 2. But a theory of quadratic forms, based on these reduced forms, may be established, which will retain at least the principal features of Gauss's theory, as may be seen from what has been developed in this paragraph.

Having established a one-to-one correspondence between certain fractions of the domain $K(\sqrt{D})$ and quadratic forms of the determinant D, we shall be able to compound fractions analogously to the well-known Gaussian composition of forms. This is readily done. Hence we may speak of species of compositions of chains analogously to that of form classes. On the researches of Gauss* we base the following: Let (a, b, c) and (a', b', c') be two quadratic forms with the same D; $a, 2b, c$ having the greatest common factor m and a', $2b', c'$ the greatest common factor m', and let m, m' be relatively prime, then the two forms may be compounded into the third form (A, B, C),

$$\frac{aa'}{d^2}=A,$$

d being the greatest common factor of $a, a', b+b'$ and

$$\frac{aB}{d}\equiv\frac{ab'}{d},\quad \frac{a'B}{d}\equiv\frac{a'b}{d},\quad \frac{b+b'}{d}B\equiv\frac{bb'+D}{d}\qquad \text{mod.}A.$$

It may be observed that the third congruence follows from the preceding two; for assume as given $b\equiv D \text{ mod.} a$, $b'\equiv D \text{ mod.} a'$ and try to determine $x^2\equiv D \text{ mod.} aa'/d$, then we obtain from $y\equiv b \text{ mod.} a$, $y\equiv b' \text{ mod.} a'/d$, $y^2\equiv D \text{ mod.} aa'/d$ the following congruences, $(y-b)(y-b')=y^2+bb'-y(b+b')\equiv 0 \text{ mod.} dA$, $(y^2+bb')/d\equiv(D+bb')/d\equiv y(b+b') \text{ mod.} A$, that is, for $y\equiv B \text{ mod.} A$ just the Gaussian formulas, and they must determine B since a/d, a'/d, $(b+b')/d$ have no common factor†. If already $(b+b')/d$ and A are

* Gauss, C. F.,: loc. cit. p. 249 and 251.

† Bachmann, I.: Grundlehren der Neueren Zahlentheorie, 1907, p. 249.

relatively prime. B will be directly obtained by the last congruence. Thus we have settled a close connection between the Gaussian classes of quadratic forms and the positively closed chains; and there also exist a very close connection between the negatively closed chains and the ideal classes. I shall in a future paper prove the following general theorem: *If* $D \equiv 2, 3 \bmod. 4$ *the number of ideal classes equals that of the negatively closed chains.*

For $D \equiv 1 \bmod. 4$ we have two types of fractions, as of quadratic forms, a well-known fact, one type characterized by odd as well as even denominators, the other by only even denominators. Hence: *If* $D \equiv 1 \bmod. 4$ *the number of ideal classes equals that of the negatively closed chains of the second type. If* $D \equiv 1 \bmod. 8$ *or* $D \equiv 5 \bmod. 8$ *with* $x^2 - Dy^2 = \pm 4$ *soluble in integers* x, y, *the number of the closed chains of either type is the same, but this is no longer true if* $D \equiv 5 \bmod. 8$, $x^2 - Dy^2 \neq \pm 4$. We verify in fact for $D = 105 \equiv 1 \bmod. 8$ four negatively reduced fractions, since $N(\eta) = -1$, $\eta = 41 + 4\sqrt{105}$, and two chains of either type, whence the number of ideal classes will be 2 with the more extensive equivalence, otherwise 4.

§ 6.

In this paragraph we shall briefly deal with chains formed in the case of negative D and shall demonstrate their connection with quadratic forms. Let us form the chain

$$(1') \qquad P_\nu^2 + D = Q_{\nu-1} Q_\nu, \qquad P_{\nu+1}^2 + D = Q_\nu Q_{\nu+1}, \ \&c.$$

or the formation of fractions

$$(1'') \qquad \frac{P_\nu + i\sqrt{D}}{Q_\nu} = \beta_\nu - \frac{P_{\nu+1} - i\sqrt{D}}{Q_\nu} = \beta_\nu - \cfrac{1}{\cfrac{P_{\nu+1} + i\sqrt{D}}{Q_{\nu+1}}},$$

where $P_{\nu+1}$ is always the absolutely least remainder of $P_\nu \bmod. Q_\nu$ and therefore $P_{\tau+1} \leq Q_\tau/2$. For $Q_\tau > D$ we have

$$P_{\tau+1}^2 + D = Q_\tau Q_{\tau+1} < (Q_\tau/2)^2 + Q_\tau < Q_\tau^2,$$

which is true with $Q_\tau > 4/3$, and hence for all $Q_\tau \geq 2$ and generally $Q_\tau > Q_{\tau+1}$. This process is to be continued till the absolutely least remainder of $P_{\rho+1}$ mod. $Q_{\rho+1}$ is $P_{\rho+1}$ itself. Then we have either $Q_\rho \geq Q_{\rho+1}$, and we shall illustrate below how we may bring this "reduced link" into connection with Gauss's reduced forms $(\pm Q_{\rho+1}, P_{\rho+1}, \pm Q_\rho)$, or we have $Q_\rho < Q_{\rho+1}$ and

$$(2) \qquad P_{\rho+1}^2 + D = Q_\rho Q_{\rho+1}, \qquad P_{\rho+2}^2 + D = Q_{\rho+1} Q_{\rho+2},$$

with $P_{\rho+2} = -P_{\rho+1}, Q_{\rho+2} = Q_\rho$, and from the latter link we take the reduced forms $(\pm Q_{\rho+2}, P_{\rho+2}, \pm Q_{\rho+1})$. For we get in both cases $(Q_{\rho+i+1}/2)^2 + D = P^2_{\rho+i+1} + D = Q_{\rho+1} Q_{\rho+i+1} > Q^2_{\rho+i+1}$, $Q_{\rho+i+1} < \sqrt{4D/3}$, $|P_{\rho+i+1}| \leq Q_{\rho+i+1}/2$, $Q_{\rho+i+1} \leq Q_{\rho+i}$ $(i = 0 \text{ or } 1)$, which are Gauss's conditions of a reduced form.*

From the reduced link we form

$$(3) \quad P^2_{\rho+2} + D = (sQ_{\rho+1} + |P_{\rho+1}|)^2 + D = Q_{\rho+1} Q_{\rho+2} = Q_{\rho+1}(s^2 Q_{\rho+1} + 2s|P_{\rho+1}| + Q_\rho),$$

with $s \gtrless 0$. For $s > 0$ we have

$$2(sQ_{\rho+1} + |P_{\rho+1}|) < s^2 Q_{\rho+1} + 2s|P_{\rho+1}| + Q_\rho,$$

$$0 < (s-1)^2 Q_{\rho+1} + 2(s-1)|P_{\rho+1}| + Q_\rho - Q_{\rho+1},$$

since $Q_\rho > Q_{\rho+1}$, and hence $|P_{\rho+2}|$ is the absolutely least remainder mod. $Q_{\rho+2}$. Then according to the rule of chain formation the next link must be formed as in (2), and after this we get once more the known reduced link. For $s = -s'$ we find by

$$0 < (s'-1)[(s'-1)Q_{\rho+1} - 2|P_{\rho+1}|] + Q_\rho - Q_{\rho+1},$$

that $P_{\rho+2}$ is still the absolutely least remainder mod. $Q_{\rho+2}$. It remains therefore to test whether in (3) any reduced link may be constructed. For $s > 0$ this is obviously impossible, but for $s = -s'$ we hav eto satisfy $s'Q_{\rho+1} - |P_{\rho+1}| \leq Q_{\rho+1}/2$, which is realized only for $s' = 1$, $|P_{\rho+1}| = Q_{\rho+1}/2$. From

$$(Q_{\rho+1}/2)^2 + D = Q_\rho Q_{\rho+1},$$
$$(Q_{\rho+1}/2)^2 + D = Q_{\rho+1} Q_\rho,$$
$$(Q_{\rho+1}/2) + D = Q_\rho Q_{\rho+1},$$

or

$$\frac{Q_{\rho+1}/2 + i\sqrt{D}}{Q_{\rho+1}} = 1 - \frac{Q_{\rho+1}/2 - i\sqrt{D}}{Q_{\rho+1}} = 1 - \cfrac{1}{\cfrac{Q_{\rho+1}/2 + i\sqrt{D}}{Q_\rho}},$$

$$\frac{Q_{\rho+1}/2 + i\sqrt{D}}{Q_\rho} = 0 - \frac{-Q_{\rho+1}/2 - i\sqrt{D}}{Q_\rho} = 0 - \cfrac{1}{\cfrac{-Q_{\rho+1}/2 + i\sqrt{D}}{Q_{\rho+1}}},$$

we have the well-known fact that the substitution of determinant $+1$, $x = x' - y'$, $y = y'$, transforms the ambiguous form $(Q_{\rho+1}, +Q_{\rho+1}/2, Q_\rho)$ into

* Gauss, C. F.: loc. cit., p. 136.

$(Q_{\rho+1}, -Q_{\rho-1}\,2, Q_\rho)$. Having $Q_{\rho-1} = Q_\rho$ in (2), we readily find the proper equivalence of the forms $(Q_\rho, +P_{\rho-1}, Q_\rho)$, $(Q_\rho, -P_{\rho+1}, Q_\rho)$, and hence the problem of the equivalence of forms with negative D is solved entirely as Gauss* has already solved it. Thus for negative D also, we can settle the problem of equivalence almost as before in the case of positive D, by the formation of chains like (1′) and (1″), and then we can also determine the number of the classes of forms.

Example. From

$$20^2+5 = 15.27,\quad 7^2+5 = 27.2,\quad 1^2+5 = 2.3,\quad 1^2+5 = 3.2,$$

or

$$\frac{20+i\sqrt{5}}{27} = 1 - \frac{7-i\sqrt{5}}{27} = 1 - \cfrac{1}{\cfrac{7+i\sqrt{5}}{2}},$$

$$\frac{7+i\sqrt{5}}{2} = 4 - \frac{1-i\sqrt{5}}{2} = 4 - \cfrac{1}{\cfrac{1+i\sqrt{5}}{3}},$$

$$\frac{1-i\sqrt{5}}{3} = 0 - \frac{1-i\sqrt{5}}{3} = 0 - \cfrac{1}{\cfrac{-1+i\sqrt{5}}{2}},$$

we find the substitution of determinant $+1$, $x = 1x' - 3y'$, $y = -1x' + 4y'$, transforms $(27x^2+40xy+15y^2)$ into the ambiguous form $(2x'^2-2x'y'+3y'^2)$. Lastly, we can give the solution of an indeterminate equation

$$x^2 - Dy^2 = A$$

in integers x and y, based on the preceding theorem in § 3 with some slight modifications, in addition to solving the more general problem of expressing a given integer A in the form (a, b, c).

* Gauss, C. F.: loc. cit. p. 138.

LUND, SWEDEN.
October, 1921.

ON CERTAIN LINEAR DIFFERENTIAL EQUATIONS OF THE SECOND ORDER.

By F. H. MURRAY.

Differential equations of the second order and of the form

$$\frac{d^2x}{dt^2} - a^2x = \mu f(x, t, \mu)$$

have been studied by Poincaré,* Bohl,† and others.

It is the purpose of this paper to generalize certain of the results obtained by Bohl for linear equations of the second order; at the same time a simplified treatment is made possible by methods of successive approximation. Among other results an analytic expression for the general integral is obtained, from which the characteristic properties of the integral curves can be immediately deduced.

While some of the results of §§ 1 and 2 are known,‡ these have been given because of their importance for later developments.§

1. In the differential equation

$$(1) \qquad \frac{d^2x}{dt^2} - \varphi(t)x = 0$$

suppose $\varphi(t)$ continuous, and positive or zero in the interval $0 \leq t \leq T$. There will be no loss of generality in restricting the analytical developments which follow to the upper right-hand quadrant of the (x, t) plane, since the transformations $x' = \pm x$, $t' = \pm t$ leave (1) unchanged or replace this equation by another of the same form.

In the interval I: $0 \leq t \leq T$ suppose

$$(2) \qquad 0 \leq \varphi(t) < B.$$

* Les Nouvelles Méthodes de la Mécanique Céleste, tome II, p. 311.

† Bulletin de la Société Mathématique de France, tome 38 (1910); see also Crelle's Journal, Band 131.

‡ See Wiman, Arkiv för Mat., Astr. och Fysik, vol. 12, Nr. 14 (1917).

§ In what follows it will be understood that discontinuous solutions are excluded from consideration.

Instead of (1), consider the equation

$$\frac{d^2 x}{d t^2} = \mu q(t) x \tag{3}$$

and assume a development of the form

$$x = x_0 + \mu x_1 + \mu^2 x_2 + \cdots$$

Suppose $x(0) = a$, $x'(0) = b$ for all values of μ.

Substituting the formal development in equation (3), we obtain the equations and boundary conditions

$$(4)\qquad \begin{aligned} &\frac{d^2 x_0}{d t^2} = 0, \qquad x_0 = a, \qquad x_0' = b \text{ if } t = 0, \\ &\frac{d^2 x_n}{d t^2} = q(t) x_{n-1}, \quad x_n = 0, \quad x_n' = 0 \text{ if } t = 0. \qquad (n = 1, 2, 3, \cdots) \end{aligned}$$

Consequently

$$(5)\qquad \begin{aligned} x_0 &= a + b t \\ x_n &= \int_0^t \int_0^z q(u) x_{n-1}(u) \, du \, dz. \qquad (n = 1, 2, 3, \cdots) \end{aligned}$$

Suppose $|x_0| < A$ in the interval I.

Then

$$|x_1| < A \int_0^t \int_0^z B \, du \, dz = A B \frac{t^2}{2},$$

and in general

$$|x_n| < A B^n \frac{t^{2n}}{(2n)!}.$$

Consequently

$$(6)\qquad |x| = \left|\sum_{n=0}^{\infty} x_n \mu^n\right| \leqq A \sum_{n=0}^{\infty} (|\mu| B)^n \frac{t^{2n}}{(2n)!}. \qquad (\text{arg } \mu)$$

It is seen immediately that for any value of μ, the series on the left of (6) can be differentiated twice term by term, and the convergence of the differentiated series is uniform with respect to t, in the interval $0 < t < T$.

Consequently the function

$$x = \sum_{n=0}^{\infty} x_n \mu^n$$

satisfies equation (3); taking $\mu = 1$ we obtain the solution of (1).

From (2), (5) it follows that so long as $a + bt > 0$, the functions $x_1(t)$, $x_2(t) \cdots\cdots$ are positive or zero; consequently so long as $x(t) > 0$, on the interval I we shall have

$$x(t) > a + bt.$$

Consider a second equation

$$\frac{d^2\bar{x}}{dt^2} - \varphi_1(t)\bar{x} = 0, \tag{7}$$

with

$$\varphi_1(t) > \varphi(t), \qquad 0 \leqq t \leqq T.$$

Suppose $\bar{x}(t)$ a solution of (7) satisfying the same initial conditions, and form the corresponding functions

$$(8) \qquad \begin{aligned} \bar{x}^0 &= a + bt \\ \bar{x}_n &= \int_0^t\int_0^z \varphi_1(u)\,\bar{x}_{n-1}(u)\,du\,dz. \end{aligned} \qquad (n = 1, 2, \cdots\cdots)$$

Since for all points of the interval $\varphi_1(u) > \varphi(u)$, we shall obtain

$$\bar{x}_1 > x_1, \cdots\cdots \bar{x}_n > x_n \tag{9}$$

since if $\bar{x}_{n-1} > x_{n-1}$, we obtain from (8) $\bar{x}_n > x_n$; applying the method of induction we obtain the inequalities (9) for all values of n under the hypothesis $a + bt > 0$.

From the equations

$$x = x_0 + x_1 + x_2 + \cdots\cdots$$
$$\bar{x} = \bar{x}_0 + \bar{x}_1 + \bar{x}_2 + \cdots\cdots$$

it follows immediately that, so long as $a + bt > 0_n$ on the interval $0 < t < T$, we shall have $\bar{x}(t) > x(t)$.

Another result to be obtained from the form of the expressions (5) is that since x_0 is linear in a, b, the same is true of x_n ($n = 1, 2, 3, \cdots$), and

$$x = aX_1 + bX_2,$$

where $X_1(t)$, $X_2(t)$ are positive or zero in the interval I. These are the *principal* functions for the value $t = 0$, since X_1 is obtained by taking $a = 1$, $b = 0$; X_2 by taking $a = 0$, $b = 1$.

We are now in a position to prove omit period.

THEOREM 1. *Assume $\varphi(t)$, $\varphi_1(t)$ continuous functions of t satisfying the conditions*

$$\varphi_1(t) > \varphi(t) \geqq 0, \qquad 0 < t \leqq T,$$

and suppose $x(t)$, $\bar{x}(t)$ solutions of the equations and boundary conditions

$$(\alpha)\qquad \begin{aligned} &\frac{d^2 x}{d t^2} - \varphi(t)\, x = 0, \qquad x = a, \qquad x' = b \text{ if } t = 0, \\ &\frac{d^2 \bar{x}}{d t^2} - \varphi_1(t)\, \bar{x} = 0, \qquad \bar{x} = a, \qquad \bar{x}' = b \text{ if } t = 0, \end{aligned}$$

in which $a > 0$; then, if $x(t) > 0$ in an interval $0 < t < t_1 < T$, $\bar{x} > x$ in the same interval.

For from equations (α) and the condition $\bar{x}'(0) = x'(0)$,

$$(10)\qquad \bar{x}'(t) - x'(t) = \int_0^t \varphi_1(\bar{x} - x)\, dt + \int_0^t x(\varphi_1 - \varphi)\, dt.$$

Since $a > 0$ the right-hand member is positive in a certain interval for which $\bar{x} > x > 0$, as was shown above; hence in a certain interval for which $\bar{x} > x > 0$, $\frac{d}{dt}(\bar{x}' - x) > 0$. But $\bar{x} - x$ can vanish a first time at $t = t_1$, only if for some preceding value $t' < t_1$, the derivative $\frac{d}{dt}(\bar{x} - x)$ vanishes. This is impossible from (10), hence the theorem.

With the aid of the principal solutions we obtain easily

THEOREM 2. *If $\varphi(t)$ is continuous, and positive or zero on the interval $0 \leqq t < T$, then if $x(t)$, $\bar{x}(t)$ are two solutions of the differential equation*

$$\frac{d^2 x}{d t^2} - \varphi(t)\, x = 0$$

which satisfy initial conditions of the form

$$\begin{aligned} x(0) &= a, \qquad & \bar{x}(0) &= \bar{a} \geqq a, \\ x'(0) &= b, \qquad & \bar{x}'(0) &= \bar{b} \geqq b, \end{aligned}$$

we shall have $\bar{x} > x$ on the interval $0 < t \leqq T$.

For in terms of the principal functions

$$\begin{aligned} x = a X_1 + b X_2, \qquad \bar{x} &= \bar{a} X_1 + \bar{b} X_2, \\ \bar{x} - x &= (\bar{a} - a) X_1 + (\bar{b} - b) X_2. \end{aligned}$$

Since

$$X_1 > 0, \qquad X_2 > 0, \qquad 0 < t \leqq T,$$

the theorem follows.

2. In the theorems above only finite intervals are considered, but the extension to infinite intervals is immediate; if φ_1, φ satisfy the condition

$$\varphi_1(t) > \varphi(t) \geqq 0$$

for all finite values of t, the conclusions of Theorems 1, 2 are valid if T has an arbitrary positive value.

For certain applications to be made it will be convenient to have

THEOREM 3. *If $\varphi(t)$ is continuous, and positive or zero for all finite values of t without vanishing identically, then the only solution of the differential equation*

$$\frac{d^2x}{dt^2} - \varphi(t)x = 0$$

which is bounded for all values of t is the solution $x = 0$.

For a value t for which $\varphi(t) > 0$ lies within a certain interval δ: $t_1 \leqq t \leqq t_2$ for which, if α is properly chosen,

$$\varphi(t) \geqq \alpha > 0. \tag{11}$$

If $x(t)$ is identically zero in δ, x must vanish for all values of t; excluding this hypothesis we can determine a sub-interval δ': $t_1' \leqq t \leqq t_2'$ for which, if β is properly chosen,

$$\pm x(t) \geqq \beta > 0. \tag{12}$$

Since x and $-x$ are bounded or unbounded simultaneously we may choose the $+$ sign in (12); with this choice

$$\varphi(t) \geqq \alpha > 0$$
$$x(t) \geqq \beta > 0 \qquad t_1' \leqq t \leqq t_2'.$$

Suppose $x'(t_1') < 0$; from the developments of the first paragraph it follows that, if $t < t_1'$,

$$x(t) \geqq x(t') + x'(t_1')(t - t_1').$$

Consequently x can have no upper bound.

Suppose $x'(t_1') > 0$; integrating the differential equation we obtain:

$$x'(t_2') - x'(t_1') = \int_{t_1}^{t_2} q(t)\, x(t)\, dt$$
$$\geqq \alpha\beta(t_2' - t_1').$$

Consequently $x'(t_2')$ must be positive, and since

$$x(t) \geqq x(t_2') + x'(t_2')(t - t_2')$$

for all finite values of $t > t_2'$, x must take on values arbitrarily large in this case also; hence the theorem.

In addition to the principal solutions we shall introduce another class of particular solutions by means of.

THEOREM 4. *If $q(t)$ is continuous, and positive or zero without vanishing identically for all finite values of t, then if (x_0, t_0) is an arbitrary set of initial values such that $x_0 \neq 0$, there exist two distinct solutions Y_1, Y_2 of the equation*

$$\frac{d^2 x}{dt^2} - q(t)\, x = 0$$

satisfying the conditions

$$Y_1(t_0) = Y_2(t_0) = x_0,$$
$$Y_1(t) < x_0, \qquad t > t_0,$$
$$Y_2(t) \leqq x_0, \qquad t < t_0.$$

By means of a transformation of the form

$$x' = \pm x,\ t' = \pm(t - t_0),$$

the discussion of the general case can be reduced to that in which $x_0 > 0$, $t_0 = 0$, $t > 0$; the transformed function $q(t')$ will continue to satisfy the conditions assumed.

Suppose $T > 0$, and determine $x(t)$ by the conditions

$$x(0) = x_0,\ x(T) = 0.$$

Since by definition the principal solutions satisfy the relations

$$X_1(0) = 1, \qquad X_1'(0) = 0,$$
$$X_2(0) = 0, \qquad X_2'(0) = 1.$$

and

$$X_1 X_2' - X_2 X_1' = C.$$

C can be determined by putting $t = 0$; hence $C = 1$.

Since

$$x = a X_1 + b X_2,$$

$$a X_1(0) + b X_2(0) = x_0,$$

$$a X_1(T) + b X_2(T) = 0,$$

$$a = x_0,$$

$$b = -x_0 \frac{X_1(T)}{X_2(T)}. \qquad (X_2(T) > 0)$$

Consider the function

$$z = -\frac{X_1(t)}{X_2(t)}.$$

(13)
$$\frac{dz}{dt} = \frac{X_1 X_2' - X_2 X_1'}{X_2^2} = \frac{1}{X_2^2}.$$

$$z(t) - z(t_1) = \int_{t_1}^{t} \frac{dt}{X_2^2}, \qquad t_1 > 0.$$

Since $X_2(0) = 0$, $X_2'(0) = 1$, it follows that, if $t > 0$, $X_2(t) \geqq t$.

Consequently

$$0 < z(t) - z(t_1) \leqq \int_{t_1}^{t} \frac{dt}{t^2}, \qquad t > t_1,$$

$$\leqq \frac{1}{t_1} - \frac{1}{t} < \frac{1}{t_1},$$

from which

(14)
$$z(t_1) < z(t) < z(t_1) + \frac{1}{t_1}, \qquad t > t_1 > 0.$$

From (13), (14) it is seen that $z(t)$ is an increasing function of t which remains less than a function of t_1; consequently z approaches a limit as t becomes infinite.

$$\lim_{t=\infty} z(t) = -\lambda.$$

Since $z < 0$, $\lambda \geqq 0$.

Suppose

$$Y_1 = x_0 [X_1 - \lambda X_2], \qquad x = x_0 [X_1 + z(T) X_2].$$

If $\bar{t} > 0$, consider the function $\bar{x}$ for $T > t$.

$$\bar{x}(\bar{t}) = x_0[X_1(\bar{t}) + z(T)X_2(\bar{t})].$$

If $t < T$, $\bar{x}'(t) < 0$; for if $\bar{x}'(t_1) \geqq 0$, $t_1 < T$, $\bar{x}(T) > \bar{x}(t_1)$. Hence for all values of $t \leqq T$,

$$\bar{x}(t) < x_0,$$

consequently

$$\bar{x}(\bar{t}) < x_0.$$

When T becomes infinite, z approaches a finite limit $-\lambda$; hence

$$\lim_{T \to \infty} \bar{x}(\bar{t}) = x_0[X_1(\bar{t}) - \lambda X_2(\bar{t})].$$

Since $\bar{x}(\bar{t}) < x_0$ for all finite values of T,

$$\lim_{T \to \infty} \bar{x}(t) = Y_1(t) \leqq x_0.$$

For negative values of t, z is positive and decreasing as $t \to -\infty$. Hence we may write

$$\lambda = \lim_{t \to -\infty} z(t), \qquad \lambda \geqq 0,$$

$$Y_2 = x_0[x_1 + \lambda X_2].$$

A discussion similar to that already given shows that, for $t < 0$, $0 < Y_2 \leqq x_0$.

The functions $Y_1(t)$, $Y_2(t)$ satisfy the conditions of the theorem if $x_0 > 0$; for if $Y_1 \equiv Y_2$, the function Y_1 would be bounded for all values of t, contrary to Theorem 3. The corresponding functions for $x_0 < 0$ are seen to be

$$Y_1 = x_0[X_1 - \lambda X_2], \qquad Y_2 = x_0[X_1 + \lambda X_2].$$

From these results it follows that there are three classes of integral curves through an arbitrary point not on the t-axis. Consider a point (x_0, t_0) in the upper half-plane, and an integral curve through this point with the tangent x_0'. If $x_0' < -\lambda x_0$, the function $x(t)$ passes from $+\infty$ to $-\infty$ when t passes from $-\infty$ to $+\infty$. If $-\lambda x_0 < x_0' < \bar{\lambda} x_0$, x passes from $+\infty$ to a certain minimum value $x_1 > 0$, and again becomes positively infinite when t passes from $-\infty$ to $+\infty$. If $x_0' > \bar{\lambda} x_0$, x passes from $-\infty$ to $+\infty$ when t passes from $-\infty$ to $+\infty$.

Since not all solutions remain bounded for $t > 0$, any two such solutions must satisfy a relation $\bar{Y}_1 = C Y_1$; similarly for $t < t_0$, $\bar{Y}_2 = C Y_2$.

The quantities $\lambda, \bar{\lambda}$ corresponding to a value $t = t_0$ have been defined by means of the principal solutions $X_1(t), X_2(t)$ corresponding to $t = t_0$;

$$
\begin{aligned}
\lambda(t_0) &= \lim_{t \to \infty} \frac{X_1(t)}{X_2(t)}, \\
-\bar{\lambda}(t_0) &= \lim_{t \to -\infty} \frac{X_1(t)}{X_2(t)}.
\end{aligned}
\tag{15}
$$

It is easily seen that if $\lambda(t_1) = 0$, then $\varphi(t) = 0$ for $t > t_1$; if $\bar{\lambda}(t_1) = 0$, $\varphi(t) = 0$ for $t < t_1$. Consequently $\lambda + \bar{\lambda} > 0$ if $\varphi(t)$ is not identically zero.

Consider the functions

$$Y_1 = x_0 [X_1 - \lambda X_2], \qquad Y_2 = x_0 [X_1 + \bar{\lambda} X_2].$$

For the value $t = t_0$,

$$\frac{Y_1'}{Y_1} = -\lambda(t_0), \qquad \frac{Y_2'}{Y_2} = \bar{\lambda}(t_0). \tag{16}$$

Corresponding to each point (t_1, x_1) of the curve $x = Y_1(t)$ the solution of (1) which remains bounded for $t > t_1$ must coincide with the solution $x = Y_1(t)$; hence for every value $t_1 > t_0$, we have the relations

$$\frac{Y_1'}{Y_1} = -\lambda(t_1), \qquad \frac{Y_2'}{Y_2} = \bar{\lambda}(t_1).$$

Since t_1 is arbitrary, these relations may be written

$$\frac{Y_1'}{Y_1} = -\lambda(t), \qquad \frac{Y_2'}{Y_2} = \bar{\lambda}(t),$$

from which

$$Y_1 = Y_1^0 e^{-\int_{t_0}^{t} \lambda\, dt}, \qquad Y_2 = Y_2^0 e^{\int_{t_0}^{t} \bar{\lambda}\, dt}. \tag{17}$$

Since the functions Y_1, Y_2 are linearly independant, it follows that *the general solution of the differential equation* (1) *can be given the form*

$$x = C_1 e^{-\int_{t_0}^{t} \lambda\, dt} + C_2 e^{\int_{t_0}^{t} \bar{\lambda}\, dt}, \tag{18}$$

in which $\lambda(t)$, $\bar{\lambda}(t)$ are positive or zero.

From (18) we deduce immediately

THEOREM 5. *A necessary and sufficient condition that each solution bounded for $t > t_0$ approach a limit different from zero as $t \to \infty$ is that a constant G exist such that*

$$\text{(a)} \qquad \int_{t_0}^{t} \lambda(t)\,dt < G, \qquad t > t_0;$$

similarly, a necessary and sufficient condition that each solution bounded for $t < t_0$ approach a limit different from zero as $t \to -\infty$ is that a constant G' exist such that

$$\int_{t}^{t_0} \lambda\,dt < G', \qquad t < t_0.$$

For condition (a) is a necessary and sufficient condition that

$$e^{-\int_{t_0}^{t} \lambda\,dt}$$

approach a limit different from zero as $t \to \infty$, and any solution bounded for $t > t_0$ can be represented in the form

$$x = Ce^{-\int_{t_0}^{t} \lambda\,dt}.$$

The second part of the theorem can be proved in a similar manner.

Suppose that $\varphi(t)$ satisfies a condition of the form

$$\varphi(t) > \alpha^2 > 0, \qquad t > t_0.$$

Applying theorem 1 to the case $\varphi_1(t) = \alpha^2$, $t > t_0$, we obtain $(\varphi(t) > \varphi_1(t))$

$$x(t) > \tfrac{1}{2}\left(x_0 + \frac{x_0'}{\alpha}\right)e^{\alpha(t-t_0)} + \tfrac{1}{2}\left(x_0 - \frac{x_0'}{\alpha}\right)e^{-\alpha(t-t_0)}, \qquad t > t_0.$$

Now suppose $x(t)$ to be the function $Y_1(t)$ corresponding to $t = t_0$; since $x(t)$ is bounded,

$$x_0 + \frac{x_0'}{\alpha} \leqq 0, \qquad \alpha \leqq -\frac{x_0'}{x_0}.$$

Since $-\dfrac{x_0'}{x_0} = \lambda(t_0)$, from (16), $\alpha \leqq \lambda(t_0)$.

If $\varphi(t) < \beta^2$, $t \geqq t_0$, we obtain in a similar manner $\lambda(t_0) \leqq \beta$.

Hence if $\varphi(t)$ is a monotone increasing function for $t > t_0$, $\alpha(t_0) = \sqrt{\varphi(t_0) - \epsilon}$, and $\lambda(t_0) > \sqrt{\varphi(t_0) - \epsilon}$; similarly if $\varphi(t)$ is a monotone decreasing function for $t > t_0$, $\varphi(t) < \varphi(t_0) + \epsilon$, and $\lambda(t_0) < \sqrt{\varphi(t_0) + \epsilon}$, for $t > t_0$. If $\varphi(t)$ approaches a limit α^2 for $t \to \infty$, $\alpha^2 - \epsilon < \varphi(t) < \alpha^2 + \epsilon$, when $t > t_1$, hence $\sqrt{\alpha^2 - \epsilon} < \lambda < \sqrt{\alpha^2 + \epsilon}$, and λ approaches α.

From Theorem 1 we obtain easily the following result:

Given the equations

$$\frac{d^2 x}{dt^2} - \varphi_1(t)\, x = 0,$$

$$\frac{d^2 x}{dt^2} - \varphi_2(t)\, x = 0,$$

in which $\varphi_2 > \varphi_1 > 0$, let $(\lambda_1, \bar{\lambda}_1)$, $(\lambda_2, \bar{\lambda}_2)$ be the pairs of functions corresponding to φ_1, φ_2 respectively; then for all values of t, $\lambda_2 > \lambda_1$, $\bar{\lambda}_2 \geqq \bar{\lambda}_1$.

For the solution of the second equation satisfying the conditions $x_2(t_0) = x_1$, $x_2'(t_0) = x_1'$ can be written in the form

$$x_2 = \frac{\bar{\lambda}_2^0 x_1 - x_1'}{\lambda_2^0 + \bar{\lambda}_2^0}\, e^{-\int_{t_0}^{t} \lambda_2\, dt} + \frac{\lambda_2^0 x_1 + x_1'}{\lambda_2^0 + \bar{\lambda}_2^0}\, e^{\int_{t_0}^{t} \bar{\lambda}_2\, dt};$$

if $-x_1' = \lambda_1^0 x_1$, $x_1 > 0$,

$$\frac{x_2}{x_1} = \frac{\bar{\lambda}_2^0 + \lambda_1^0}{\lambda_2^0 + \bar{\lambda}_2^0}\, e^{-\int_{t_0}^{t} \lambda_2\, dt} + \frac{\lambda_2^0 - \lambda_1^0}{\lambda_2^0 + \bar{\lambda}_2^0}\, e^{\int_{t_0}^{t} \bar{\lambda}_2\, dt}.$$

Since* $x_1(t) \to 0$, and $x_1 < x_2$, by Theorem 1, x_2 must remain positive; since the second integral on the right can not remain bounded, $\lambda_2^0 \geqq \lambda_1^0$. Similarly if $x_1' = \bar{\lambda}_1 x_1$,

$$\frac{x_2}{x_1} = \frac{\bar{\lambda}_2^0 - \bar{\lambda}_1^0}{\lambda_2^0 + \bar{\lambda}_2^0}\, e^{-\int_{t_0}^{t} \lambda_2\, dt} + \frac{\lambda_2^0 + \bar{\lambda}_1^0}{\lambda_2^0 + \bar{\lambda}_2^0}\, e^{\int_{t_0}^{t} \bar{\lambda}_2\, dt}.$$

Here again x_2 must remain positive, and $\bar{\lambda}_2^0 \geqq \bar{\lambda}_1^0$. Since t_0 is an arbitrary value of t the result stated follows.

More definite results concerning the behavior of the solutions as $t \to \infty$ can be obtained if $\varphi(t)$ satisfies a condition of the form

$$\alpha^2 t^{m_1} < \varphi(t) < \alpha^2 t^{m_2}, \qquad t > t_0.$$

$$\text{(a)} \qquad -2 < m_1 < m_2.$$

* The function $x_1(t)$ is assumed to satisfy the first differential equation above, and $x_1(t_0) = x_1$, $x_1'(t_0) = x_1'$.

or

(b) $$m_1 < m_2 < -2.$$

For if $m \neq -2$, the equation

(c) $$\frac{d^2 y}{dt^2} - a^2 t^m y = 0$$

can be transformed into the equation*

(d) $$\frac{d^2 y}{dz^2} + \left[-\tfrac{1}{4} + \frac{\frac{1}{4} - p^2}{z^2}\right] y = 0,$$

by means of the substitutions

(e) $$p = \frac{1}{m+2}, \qquad y = t^{-\frac{m}{4}} y, \qquad z = \frac{4a}{m+2} t^{\frac{m+2}{2}}.$$

If condition (a) is satisfied, $z \to \infty$, as $t \to \infty$, and equation (d) has a solution $y = W_{0p}(z)$ which approaches zero as $z \to \infty$:

$$W_{0p}(z) = e^{-\frac{z}{2}}\left[1 + \frac{a_1}{z} + \cdots + \frac{a_n}{z^n} + \frac{R_n}{z^{n+1}}\right],$$

$$|R_n(z)| < \epsilon, \qquad z > z_0.$$

The corresponding solution of (c) can be written

$$y = t^{-\frac{m}{4}} e^{-\frac{2a}{m+2} t^{\frac{m+2}{2}}}\left[1 + R_1(t)\, t^{-\frac{m+2}{2}}\right].$$

Giving m the values m_1, m_2, we obtain the result that if $(\lambda_1, \bar{\lambda}_1)$, $(\lambda_2, \bar{\lambda}_2)$, $(\lambda, \bar{\lambda})$ are the functions corresponding to equation (c) and the equation (1) respectively,

$$\lambda_2 > \lambda \geqq \lambda_1, \qquad \bar{\lambda}_2 \geqq \bar{\lambda} > \bar{\lambda}_1.$$

Then if

$$y_1 = x_0 e^{-\int_{t_0}^{t} \lambda_1 dt}, \qquad x = x_0 e^{-\int_{t_0}^{t} \lambda dt}, \qquad y_2 = x_0 e^{-\int_{t_0}^{t} \lambda_2 dt},$$

$$y_2 \leqq x \leqq y_1, \qquad t > t_0.$$

---

* See Whittaker and Watson, Modern Analysis, Chap. XVI.

from which

$$C_2' t^{\frac{-m_2}{4}} e^{\frac{-2a}{m_2+2} t^{\frac{m_2+2}{2}}} \left[1 + R_2 t^{\frac{-m_2+2}{2}}\right]$$

$$\leq x \leq C_1' t^{\frac{-m_1}{4}} e^{\frac{-2a}{m_1+2} t^{\frac{m_1+2}{2}}} \left[1 + R_1 t^{\frac{-m_1+2}{2}}\right].$$

In case (b) $z \to 0$ as $t \to \infty$; the solutions of the transformed equation can be given in terms of the functions $M_{0p}(z)$, $M_{0,-p}(z)$ where

$$M_{0p}(z) = z^{\frac{1}{2}+p}\left[1 + \sum_{n=1}^{\infty} a_n z^{2n}\right].$$

From (e),

$$y = C\left[1 + \sum_{n=1}^{\infty} a_n' t^{n(m+2)}\right].$$

Consequently in case (b) the solutions of (c) which remain bounded approach limits different from zero as $t \to \infty$, and the same must be true of the corresponding solutions of (1); for, as in case (a), we obtain inequalities $y_1 \leqq x \leqq y_2$, where $y_1(t_0) = y_2(t_0) = x(t_0)$, y_1, y_2 satisfying equations of type (c), and both y_1 and y_2 approach positive limits as $t \to \infty$.

Results concerning the manner in which solutions of (1) become infinite as $t \to \infty$ can be easily obtained from the comparison theorems above in cases (a) and (b), but these will not be developed here.

Similar remarks apply to the function $\lambda(t)$, for $t < t_0$.

3. The results of the first two paragraphs can be extended to the more general equation

$$\frac{d}{dt}\left[k(t)\frac{dx}{dt}\right] - \varphi(t)x = 0, \tag{19}$$

if $k(t)$ is continuous and satisfies a relation of the form

$$0 < a \leq k(t) \leq b$$

for all real, finite values of t. For this equation can be written

$$k(t)\frac{d}{dt}\left[k(t)\frac{dx}{dt}\right] - k(t)\varphi(t)x = 0. \tag{19'}$$

Suppose

$$u(t) = \int_0^t \frac{dt}{k(t)}.$$

The relation between u and t is one-to-one and continuous; if

$$\varphi(u) = k(t)\varphi(t),$$

equation (19′) can be given the form

$$(20) \qquad \frac{d^2 x}{d u^2} - \varphi(u) x = 0.$$

Since this equation is of the type (1), the results already obtained can be applied.

4. In this paragraph will be given a qualitative study of the solutions of the non-homogeneous equation

$$(21) \qquad \frac{d^2 x}{d t^2} - \varphi(t) x = \psi(t)$$

in which φ, ψ are continuous for all real, finite values of t, and satisfy conditions of the form

$$(22) \qquad 0 < \beta^2 < \varphi(t) < \alpha^2, \qquad |\psi(t)| < A.$$

Equation (21) can be given the form

$$\frac{d^2 x}{d t^2} - \alpha^2 x = [\varphi(t) - \alpha^2] x + \psi(t).$$

If $f(t) = \varphi(t) - \alpha^2$, then $|f| < \alpha^2 - \beta^2 = \gamma$. Replace $f(t)$ by $\mu f(t)$:

$$(23) \qquad \frac{d^2 x}{d t^2} - \alpha^2 x = \mu f(t) x + \psi(t).$$

Assume a development of the form

$$(24) \qquad x = x_0 + \mu x_1 + \cdots + \mu^n x_n + \cdots.$$

From (23)

$$(25) \qquad \begin{aligned} &\frac{d^2 x_0}{d t^2} - \alpha^2 x_0 = \psi(t), \\ &\frac{d^2 x_n}{d t^2} - \alpha^2 x_n = f(t) x_{n-1}. \qquad (n = 1, 2, 3, \cdots) \end{aligned}$$

The general solution of the first equation can be obtained in the form

$$x_0 = \frac{1}{2\alpha}\left[e^{\alpha t}\int_{c_1}^{t} e^{-\alpha u}\psi(u)\,du - e^{-\alpha t}\int_{c_2}^{t} e^{\alpha u}\psi(u)\,du\right].$$

Suppose $\alpha > 0$; a particular solution can be obtained by choosing $C_1 = \infty$, $C_2 = -\infty$, since the infinite integrals converge,

$$x_0 = \frac{-1}{2\alpha}\left[e^{\alpha t}\int_t^{\infty} e^{-\alpha u}\psi(u)\,du + e^{-\alpha t}\int_{-\infty}^{t} e^{\alpha u}\psi(u)\,du\right]. \tag{26}$$

From (22)

$$|x_0| < \frac{A}{2\alpha}\left[e^{\alpha t}\int_t^{\infty} e^{-\alpha u}\,du + e^{-\alpha t}\int_{-\infty}^{t} e^{\alpha u}\,du\right] = \frac{A}{\alpha^2}.$$

Similarly, we may choose

$$x_n = \frac{-1}{2\alpha}\left[e^{\alpha t}\int_t^{\infty} e^{-\alpha u} f(u) x_{n-1}(u)\,du + e^{-\alpha t}\int_{-\infty}^{t} e^{\alpha u} f(u) x_{n-1}(u)\,du\right]. \tag{27}$$

If $|x_{n-1}| < A_{n-1}$, then $|x_n| < \frac{\gamma}{\alpha^2} A_{n-1} = A_n$. Hence $|x_0| < \frac{A}{\alpha^2}$.

$$|x_n| < \frac{A\gamma^n}{\alpha^{2n+2}}. \qquad (n = 1, 2, 3, \cdots) \tag{28}$$

The series (24) is dominated by the series

$$\frac{A}{\alpha^2}\left[1 + \frac{\mu\gamma}{\alpha^2} + \left(\frac{\mu\gamma}{\alpha^2}\right)^2 + \cdots + \left(\frac{\mu\gamma}{\alpha^2}\right)^n + \cdots\right]. \tag{29}$$

and must therefore converge uniformly for all real values of t when $\mu = 1$, since $\gamma < \alpha^2$. The function $x(t)$ has the upper bound

$$\frac{A}{\alpha^2}\,\frac{1}{1 - \frac{\gamma}{\alpha^2}} = \frac{A}{\beta^2}.$$

The differentiated series also converges uniformly, consequently the function $x(t)$ is a particular solution of equation (21) which remains bounded for all real values of t. There can be only one such solution, since the difference of two particular solutions satisfies the homogeneous equation

$$\frac{d^2x}{dt^2} - \varphi(t)\,x = 0,$$

to which Theorem 3 can be applied. We have thus proved

THEOREM 6. *If $\varphi(t)$, $\psi(t)$ are continuous and satisfy conditions of the form*

$$0<\beta^2<\varphi(t)<\alpha^2, \quad |\psi(t)|<A,$$

for all real, finite values of t, then there is one and only one solution of the differential equation

$$\frac{d^2x}{dt^2}-\varphi(t)\,x=\psi(t),$$

which remains bounded for all real values of t.

The method of successive approximation employed in the demonstration of Theorem 6 is convenient for later developments; but with the aid of the results of § 2 it can be shown that *under the conditions $0<\beta^2<\varphi(t)$, $|\psi(t)|<A$, there exists just one solution of the equation*

$$\frac{d^2x}{dt^2}-\varphi(t)\,x=\psi(t),$$

which remains bounded for all real values of t.

For suppose

$$y_1=e^{-\int_{t_0}^{t}\lambda\,dt}, \qquad y_2=e^{\int_{t_0}^{t}\bar{\lambda}\,dt}.$$

Then

$$y_1y_2'-y_2y_1'=(\bar{\lambda}+\lambda)\,e^{\int_{t_0}^{t}(\bar{\lambda}-\lambda)\,dt}=C,$$

$$=\lambda(t_0)+\bar{\lambda}(t_0).$$

Consequently

$$\text{(a)} \qquad \frac{e^{-\int_{t_0}^{t}\lambda\,dt}}{\lambda(t)+\bar{\lambda}(t)}=\frac{e^{\int_{t_0}^{t}\bar{\lambda}\,dt}}{\bar{\lambda}(t_0)+\lambda(t_0)}.$$

If t_1, t_2 are arbitrary values of t, a solution of the non-homogeneous equation can be given the form

$$x=\frac{-1}{\lambda(t_0)+\bar{\lambda}(t_0)}\left[\int_{t_1}^{t}\psi(z)\,y_1(t)\,y_2(z)\,dz+\int_{t}^{t_2}\psi(z)\,y_2(t)\,y_1(z)\,dz\right]$$

$$=\frac{-1}{\lambda(t_0)+\bar{\lambda}(t_0)}\left[\int_{t_1}^{t}\psi(z)\,e^{-\int_{t_0}^{t}\lambda\,du+\int_{t_0}^{z}\bar{\lambda}\,du}\,dz+\int_{t}^{t_2}\psi(z)\,e^{-\int_{t_0}^{z}\lambda\,du+\int_{t_0}^{t}\bar{\lambda}\,du}\,dz\right].$$

From (a) and the corresponding equation with t replaced by z, we obtain

$$-x = \int_{t_1}^{t} \frac{\psi(z)e^{-\int_z^t \lambda du}}{\lambda(z)+\bar{\lambda}(z)}dz + \int_{t}^{t_2} \frac{\psi(z)e^{-\int_t^z \bar{\lambda} du}}{\lambda(t)+\bar{\lambda}(t)}dz.$$

Since $\varphi - \beta^2 > 0$, it follows from the results of § 2 that $\lambda \geqq \beta$, $\bar{\lambda} \geqq \beta$. Consequently

$$\begin{aligned} |x| &\leqq \frac{A}{2\beta}\left[\int_{t_1}^{t} e^{-\int_z^t \beta du} dz + \int_{t}^{t_2} e^{-\int_t^z \beta du} dz\right] \\ &\leqq \frac{A}{2\beta^2}\left[2 - e^{-\beta(t-t_1)} - e^{\beta(t-t_2)}\right]. \end{aligned}$$

Hence the integrals defining x converge if $t_1 = -\infty$, $t_2 = +\infty$, and $|x| < \frac{A}{\beta^2}$. The function

$$x = -\int_{-\infty}^{t} \frac{\psi(z)e^{\int_z^t \lambda du}}{\lambda(z)+\bar{\lambda}(z)}dz - \int_{t}^{\infty} \frac{\psi(z)e^{\int_t^z \bar{\lambda} du}}{\lambda(t)+\bar{\lambda}(t)}dz$$

is one solution of the non-homogeneous equation which remains bounded, and by Theorem 3 there can be only one such solution.

By means of the transformation of § 3 these results can be extended to the equation

$$\frac{d}{dt}\left[k(t)\frac{dx}{dt}\right] - \varphi(t)x = \psi(t), \tag{30}$$

if (30) is written in the form

$$\frac{d^2x}{du^2} - \Phi(u)x = \Psi(u),$$

$$\Phi(u) = k(t)\varphi(t), \qquad \Psi(u) = k(t)\psi(t).$$

In particular the series expansion of which the general term is given by (27) can be employed.

From the form of (27) the sign of $x(t)$ is constant if $\psi(t) \not\equiv 0$. For suppose $\psi(t) \geqq 0$ without vanishing identically. Then from (26), $x_0 < 0$. Since $f(t)$ is negative, x_n has the sign of x_{n-1}, hence $x_1 < 0$, $x_2 < 0$, $\cdots\cdots$; consequently $x < 0$ for all real values of t. Similarly if $\psi \leqq 0$, $x > 0$.

5. Differential equations of the type considered here are especially important when the coefficients φ, ψ are quasi-periodic* functions of t.

By definition $f(t)$ is a uniformly continuous quasi-periodic function of t with the periods $\alpha_1, \alpha_2, \cdots \alpha_m$ if, given $\varepsilon > 0$, an η can be determined such that

$$|f(t+\tau)-f(t)| < \varepsilon$$

for all real values of t when τ satisfies the conditions

$$\left|\frac{\tau}{\alpha_i} - n_i\right| < \eta, \qquad (i = 1, 2, \cdots m)$$

$n_1 \cdots n_m$ being integers. In the works of Bohl it is shown that a function satisfying these conditions is necessarily bounded.

It is proposed to demonstrate the following

THEOREM 7. *If $k(t)$, $\varphi(t)$, $\psi(t)$ are uniformly continuous quasi periodic functions of t with the periods $\alpha_1 \cdots \alpha_m$, and such that*

$$0 < b^2 < k(t) < a^2, \quad 0 < \beta^2 < \varphi(t) < \alpha^2, \quad |\psi(t)| < A$$

$$\int \frac{dt}{k(t)} = ct + g(t),$$

$g(t)$ being a uniformly continuous quasi-periodic function, then there exists just one solution of the differential equation

$$\frac{d}{dt}\left[k(t)\frac{dx}{dt}\right] - \varphi(t)x = \psi(t),$$

which is bounded for all real values of t; this solution is quasi-periodic with the periods $\alpha_1, \alpha_2 \cdots \alpha_m$.

It was shown by Bohl that $g(t)$ is quasi-periodic with the same periods $\alpha_1 \cdots \alpha_m$.

Taking

$$z = \int_0^t \frac{dt}{k(t)} = ct + g(t), \qquad c > 0,$$

$$\Phi(z) = k(t)\varphi(t), \qquad \Psi(z) = k(t)\psi(t),$$

* See Esclangon, Nouvelles Recherches sur les fonctions quasi-periodiques. Annales de l'Observatoire de Bordeaux, XVI, (1917).

we obtain the differential equation

$$\frac{d^2 x}{d z^2} - \Phi(z) x = \Psi(z). \tag{31}$$

It has already been seen that there exists just one bounded solution of this equation: a uniformly convergent series expansion of this solution can be obtained from (26), (27), replacing t by z. It will be shown that each term of this series is quasi-periodic with the periods $\alpha_1 \cdots \alpha_m$.

From (26)

$$x_0 = \frac{-1}{2\alpha}\left[\int_z^{\infty} e^{-\alpha(u-z)} \Psi(u)\, du + \int_{-\infty}^{z} e^{-\alpha(z-u)} \Psi(u)\, du\right].$$

Substituting

$$u = \int_0^v \frac{dv}{k(v)},$$

then

$$x_0 = \frac{-1}{2\alpha}(I_1 + I_2),$$

where

$$I_1 = \int_z^{\infty} e^{-\alpha(u-z)} \Psi(u)\, du = e^{\alpha g(t)} \int_t^{\infty} e^{-\alpha c(v-t)} [e^{-\alpha g(v)} \psi(v)]\, dv.$$

If $g(t)$ is quasi-periodic, $e^{-\alpha g(t)}$ is quasi-periodic with the same periods; since the product of two quasi-periodic functions is quasi-periodic, it will be sufficient to show that, if $P(t)$ is a uniformly continuous quasi-periodic function with the periods $\alpha_1 \cdots \alpha_m$, the same is true of

$$Q(t) = \int_t^{\infty} e^{-\alpha c(v-t)} P(v)\, dv.$$

In the integral,

$$Q(t+\tau) = \int_{t+\tau}^{\infty} e^{-\alpha c(v-t-\tau)} P(v)\, dv,$$

substitute $v = w + \tau$:

$$Q(t+\tau) = \int_t^{\infty} e^{-\alpha c(w-t)} P(w+\tau)\, dw,$$

$$\begin{aligned} Q(t+\tau) - Q(t) &= \int_t^{\infty} e^{-\alpha c(w-t)} [P(w+\tau) - P(w)]\, dw \\ &= [P(\zeta+\tau) - P(\zeta)] \int_t^{\infty} e^{-\alpha c(w-t)}\, dw \\ &= \frac{1}{\alpha c} [P(\zeta+\tau) - P(\zeta)]. \qquad (\zeta > t) \end{aligned}$$

Hence if τ is so chosen that for all values of t, $|P(t+\tau)-P(t)|<\epsilon$, then $|Q(t+\tau)-Q(t)|<\frac{\epsilon}{\alpha c}$.

It follows that $Q(t)$ is a uniformly continuous quasi-periodic function with the same periods, and the same is true of $I_1(t)$. An almost identical discussion shows that $I_2(t)$ is likewise quasi-periodic.

Under the assumption that $x_{n-1}(t)$ and $f(t)$ are uniformly continuous quasi-periodic functions with the given periods, ($\bar{f}(t)$ corresponding to $f(t)$ in equation (27)), the same discussion shows that $x_n(t)$ satisfies these conditions also. Hence by the principle of induction each term of the uniformly convergent series

$$S(t)=x_0+x_1+\cdots+x_n+\cdots$$

satisfies these conditions, from which it follows immediately that the function $S(t)$ is a uniformly continuous quasi-periodic function with the periods $\alpha_1\cdots\alpha_m$.

Making use of the fact that $k(t)>0$, we obtain the result that *the general solution of the equation*

$$\frac{d}{dt}\left[k(t)\frac{dx}{dt}\right]-q(t)x=\psi(t)$$

can be given the form

$$x=C_1e^{-\int\lambda dt}+C_2e^{\int\bar{\lambda}dt}+S(t),$$

in which $\lambda(t)$, $\bar{\lambda}(t)$ are positive or zero, and $S(t)$ is quasi-periodic with the periods $\alpha_1\cdots\alpha_m$.

SPHERICAL REPRESENTATION OF CONJUGATE SYSTEMS AND ASYMPTOTIC LINES.*

BY W. C. GRAUSTEIN.

1. **Introduction.** A system of curves on the Gauss sphere represents, as is well known, a conjugate system of curves on each of infinitely many surfaces. It is shown in this paper that, if the ratio of the radii of normal curvature in the conjugate directions is prescribed subject to a certain condition, *one* of the required surfaces is determined to within its homothetics (§ 2).

Since the condition in question can be written in a form involving only the differences of the point and of the plane invariants of the conjugate system and an expression whose vanishing is the condition that the conjugate system be isothermal-conjugate, the general result is particularly adaptable to the important special cases of conjugate systems which have equal invariants of either kind or are isothermal-conjugate (§§ 3, 4). It is also readily applied to translation surfaces (§ 5).

In the case of lines of curvature the condition can be put into a second equally striking form containing, besides the differences of the point and of the plane invariants, merely an expression whose vanishing is the condition that the ratio of the principal radii of curvature is of the form $U(u)/V(v)$, the lines of curvature being parametric. Inasmuch as the lines of curvature, or their spherical representation, form an isothermal (orthogonal) system, according as the point invariants, or the plane invariants, are equal, the general result again leads immediately to important special theorems (§ 6).

One family of a system of curves on the Gauss sphere represents one family of asymptotic lines on each of infinitely many surfaces having the prescribed spherical representation. In § 7 it is proved that, if the second family of asymptotic lines is prescribed subject to a certain condition, *one* of the surfaces is determined to within its homothetics. Thus a generalization of Dini's Theorem† is obtained. Applications of it are made to ruled and minimal surfaces.

The content of the paper is closely related to the theory of parallel maps and can, in fact, be used as the basis of a development of this theory. For, a parallel map can be considered as a one-to-one point correspondence between

---

* Presented to the American Mathematical Society, February 25, 1922.

† Cf., e. g., Eisenhart, Differential Geometry, p. 192.

two surfaces such that their spherical representations are identical; furthermore, in the theory of a general parallel map the corresponding conjugate systems and the quotient of the ratios of the radii of normal curvature in the corresponding conjugate directions play determinative roles.*

It is assumed throughout that the surfaces considered are real, non-developable, and analytic. The families of curves constituting a conjugate system can be taken as real or, provided the minimal curves of a minimal surface are excluded,† as conjugate-imaginary.

2. **Conjugate system arbitrary.** If a surface S, $x = x(u, v)$:

$$x_1 = x_1(u, v), \qquad x_2 = x_2(u, v), \qquad x_3 = x_3(u, v),$$

is referred to a conjugate system as the parametric curves, the Codazzi equations can be written in the form

$$(1) \qquad \frac{\partial \log e}{\partial v} = \begin{Bmatrix}12\\1\end{Bmatrix}' - \begin{Bmatrix}11\\2\end{Bmatrix}' \frac{g}{e}, \qquad \frac{\partial \log g}{\partial u} = \begin{Bmatrix}12\\2\end{Bmatrix}' - \begin{Bmatrix}22\\1\end{Bmatrix}' \frac{e}{g},$$

where e, $f (= 0)$, g are the differential coefficients of the second order of S and the Christoffel symbols pertain to the linear element,

$$\mathfrak{E}\, du^2 + 2\mathfrak{F}\, du\, dv + \mathfrak{G}\, dv^2,$$

of the spherical representation of S. From equations (1) we obtain the relation

$$(2) \qquad \frac{\partial^2 \log e/g}{\partial u\, \partial v} = \frac{\partial}{\partial u}\left[\begin{Bmatrix}12\\1\end{Bmatrix}' - \begin{Bmatrix}11\\2\end{Bmatrix}' \frac{g}{e}\right] - \frac{\partial}{\partial v}\left[\begin{Bmatrix}12\\2\end{Bmatrix}' - \begin{Bmatrix}22\\1\end{Bmatrix}' \frac{e}{g}\right].$$

Since

$$(3) \qquad E = \frac{e^2 \mathfrak{G}}{\mathfrak{H}^2}, \qquad F = -\frac{eg\mathfrak{F}}{\mathfrak{H}^2}, \qquad G = \frac{g^2 \mathfrak{E}}{\mathfrak{H}^2},$$

---

* Cf. Author, "Parallel maps of surfaces", Trans. Amer. Math. Soc., vol. 23 (1922), pp. 298–332, in particular §§ 19, 20, where the results of the present paper are used in developing conditions for the existence of a parallel map when the spherical representation, together with certain other definitive elements, is given. It is to be noted that the theory of the present § 7, concerning asymptotic lines, comes into play in the case of a parallel map for which the corresponding conjugate systems have become single families of asymtotic lines.

† The ratio of the radii of normal curvature in a pair of minimal conjugate directions is undefined.

where $\mathfrak{H}^2 = \mathfrak{E}\mathfrak{G} - \mathfrak{F}^2$, it follows that

$$(4) \qquad \frac{e}{g} = \frac{\mathfrak{E}}{\mathfrak{G}} \frac{R_1}{R_2},$$

where R_1 and R_2 are the radii of normal curvature of S in the given conjugate directions. Hence the ratio R_1/R_2 satisfies the equation:

$$(5) \qquad \frac{\partial^2}{\partial u\, \partial v} \log \frac{\mathfrak{E}}{\mathfrak{G}} \frac{R_1}{R_2} = \frac{\partial}{\partial u}\left[\begin{Bmatrix}12\\1\end{Bmatrix}' - \frac{\mathfrak{G}}{\mathfrak{E}}\begin{Bmatrix}11\\2\end{Bmatrix}' \frac{R_2}{R_1}\right] - \frac{\partial}{\partial v}\left[\begin{Bmatrix}12\\2\end{Bmatrix}' - \frac{\mathfrak{E}}{\mathfrak{G}}\begin{Bmatrix}22\\1\end{Bmatrix}' \frac{R_1}{R_2}\right].$$

Conversely, if there is given on the Gauss sphere a system of curves $\mathfrak{C}$ and a point function R_1/R_2 such that, when the curves $\mathfrak{C}$ are parametric*, $\mathfrak{E}$, $\mathfrak{F}$, $\mathfrak{G}$ and R_1/R_2 satisfy (5), equations (1) and (4), in e and g, are compatible; e and g can be found by a quadrature and are unique except for the same multiplicative constant, k. Consequently, the surface S is determined to within its homothetics.

To find the parametric equations, $x = x\,(u,\, v)$, of S, it is necessary first to find those of the sphere, $\zeta = \zeta\,(u,\, v)$, by solving a Riccati equation. Then x_1, x_2, x_3 can be obtained by quadratures from the equations

$$\frac{\partial x}{\partial u} = \frac{e}{\mathfrak{H}^2}\left(-\mathfrak{G}\frac{\partial \zeta}{\partial u} + \mathfrak{F}\frac{\partial \zeta}{\partial v}\right), \qquad \frac{\partial x}{\partial v} = \frac{g}{\mathfrak{H}^2}\left(\mathfrak{F}\frac{\partial \zeta}{\partial u} - \mathfrak{E}\frac{\partial \zeta}{\partial v}\right),$$

and are determined except for the multiplier k and additive constants.

THEOREM 1. *A system of curves on the sphere represents a conjugate system on a surface for which the ratio R_1/R_2 of the radii of normal curvature in the conjugate directions is prescribed if and only if $\mathfrak{E}$, $\mathfrak{F}$, $\mathfrak{G}$, R_1/R_2 satisfy* (5). *The surface is then determined to within its homothetics and its point coordinates can be found by quadratures, when those of the sphere are known.*

Since between the Christoffel symbols for S and those for the spherical representation the relations

$$(6) \qquad \begin{Bmatrix}12\\1\end{Bmatrix} = -\frac{g}{e}\begin{Bmatrix}11\\2\end{Bmatrix}', \qquad \begin{Bmatrix}12\\2\end{Bmatrix} = -\frac{e}{g}\begin{Bmatrix}22\\1\end{Bmatrix}'$$

subsist, (2) can be written in the form

$$(7) \qquad \frac{\partial^2 \log e/g}{\partial u\, \partial v} = (h-k) + (h'-k'),$$

* Explicit mention of this condition, which we shall always assume fulfilled, will henceforth be suppressed.

where

$$(8)\qquad h-k=\frac{\partial}{\partial u}\left\{\begin{matrix}12\\1\end{matrix}\right\}-\frac{\partial}{\partial v}\left\{\begin{matrix}12\\2\end{matrix}\right\},\qquad h'-k'=\frac{\partial}{\partial u}\left\{\begin{matrix}12\\1\end{matrix}\right\}'-\frac{\partial}{\partial v}\left\{\begin{matrix}12\\2\end{matrix}\right\}',$$

and h, k are the point invariants and h', k' the plane invariants of the given conjugate system.

3. Isothermal-conjugate systems. If the parametric curves on the surface S form an isothermal-conjugate system, e/g is of the form $U(u)/V(v)$ and (5) splits into the two equations:

$$(9)\qquad \begin{gathered}\frac{\partial^2}{\partial u\,\partial v}\log\frac{\mathfrak{E}}{\mathfrak{G}}\frac{R_1}{R_2}=0,\\ \frac{\partial}{\partial u}\left[\left\{\begin{matrix}12\\1\end{matrix}\right\}'-\frac{\mathfrak{G}}{\mathfrak{E}}\left\{\begin{matrix}11\\2\end{matrix}\right\}'\frac{R_2}{R_1}\right]=\frac{\partial}{\partial v}\left[\left\{\begin{matrix}12\\2\end{matrix}\right\}'-\frac{\mathfrak{E}}{\mathfrak{G}}\left\{\begin{matrix}22\\1\end{matrix}\right\}'\frac{R_1}{R_2}\right].\end{gathered}$$

THEOREM 2. *A necessary and sufficient condition that there exist a surface S having an isothermal-conjugate system which is represented by a given system of curves on the sphere is that equations (9), in R_1/R_2, be compatible. Then to each solution, R_1/R_2, there corresponds a surface S which is unique to within its homothetics and the ratio of whose radii of normal curvature in the isothermal-conjugate directions is R_1/R_2.*

If the given system of curves on the sphere is to represent an isothermal-conjugate system and the parameters u, v are to be isothermal-conjugate as well, that is, if e/g is to have the value $+1$ or -1, according as S is to be of positive or negative curvature, equations (9) reduce to the single, well known condition:*

$$\frac{\partial}{\partial u}\left[\left\{\begin{matrix}12\\1\end{matrix}\right\}'+\left\{\begin{matrix}11\\2\end{matrix}\right\}'\right]=\frac{\partial}{\partial v}\left[\left\{\begin{matrix}12\\2\end{matrix}\right\}'+\left\{\begin{matrix}22\\1\end{matrix}\right\}'\right].$$

If this condition is satisfied, S is unique to within its homothetics and $R_1/R_2=\pm\mathfrak{G}/\mathfrak{E}$.

From (7) we conclude the following:

THEOREM 3. *A necessary and sufficient condition that a conjugate system be isothermal-conjugate is that the difference of its point invariants equal the difference of its plane invariants taken in the opposite order.*

4. Conjugate systems with equal invariants. By using (7), in conjunction with (4) and (6), we obtain from Theorem 1 the following results.

* Cf., e. g., Eisenhart, Differential Geometry, p. 202.

THEOREM 4. *A necessary and sufficient condition that there exist a surface S having a conjugate system with equal point invariants which is represented by a given system of curves on the sphere is that the equations*

$$(10)\quad \frac{\partial}{\partial u}\left[\frac{\mathfrak{G}}{\mathfrak{E}}\begin{Bmatrix}11\\2\end{Bmatrix}'\frac{R_2}{R_1}\right]=\frac{\partial}{\partial v}\left[\frac{\mathfrak{E}}{\mathfrak{G}}\begin{Bmatrix}22\\1\end{Bmatrix}'\frac{R_1}{R_2}\right],\quad \frac{\partial^2}{\partial u\,\partial v}\log\frac{\mathfrak{E}}{\mathfrak{G}}\frac{R_1}{R_2}=h'-k',$$

in R_1/R_2, be compatible. Then to each solution, R_1/R_2, there corresponds a surface S which is unique to within its homothetics and has R_1/R_2 as the ratio of its radii of normal curvature in the conjugate directions.

The given curves on the sphere represent a conjugate system with equal plane invariants on some surface S if and only if $h'=k'$; then each solution of

$$\frac{\partial^2}{\partial u\,\partial v}\log\frac{\mathfrak{E}}{\mathfrak{G}}\frac{R_1}{R_2}=\frac{\partial}{\partial v}\left[\frac{\mathfrak{E}}{\mathfrak{G}}\begin{Bmatrix}22\\1\end{Bmatrix}'\frac{R_1}{R_2}\right]-\frac{\partial}{\partial u}\left[\frac{\mathfrak{G}}{\mathfrak{E}}\begin{Bmatrix}11\\2\end{Bmatrix}'\frac{R_2}{R_1}\right]$$

leads to a surface S with the desired properties.

In order that the conjugate system have both equal point invariants and equal plane invariants, it is necessary and sufficient that $h'=k'$ and that the first equations in (9) and (10) be compatible in R_1/R_2. In this connection we note, from (7), the following known result:

THEOREM 5. *If a conjugate system has two of the three properties, (a) of having equal point invariants, (b) of having equal plane invariants, (c) of being isothermal-conjugate, it has also the third.*

5. **Translation surfaces.** A surface is a translation surface referred to its generators as the parametric curves if and only if

$$\begin{Bmatrix}12\\1\end{Bmatrix}=0,\qquad \begin{Bmatrix}12\\2\end{Bmatrix}=0.$$

Accordingly, we obtain, from (6) and (5), the following result.

THEOREM 6. *A necessary and sufficient condition that a system of curves on the sphere represent the generators of a translation surface is that*

$$\begin{Bmatrix}11\\2\end{Bmatrix}'=0,\qquad \begin{Bmatrix}22\\1\end{Bmatrix}'=0.$$

Then each solution, R_1/R_2, of the second equation of (10) *determines such a surface to within its homothetics and the ratio of the radii of normal curvature in the directions of the generators is R_1/R_2.*

6. **Lines of curvature.** If the given conjugate system on the surface S consists of the lines of curvature, we obtain, from (3) and (4), the equations

$$(11)\qquad \frac{e}{g} = \frac{\mathfrak{E}}{\mathfrak{G}}\,\frac{r_1}{r_2}, \qquad \frac{E}{G} = \frac{\mathfrak{E}}{\mathfrak{G}}\left(\frac{r_1}{r_2}\right)^2,$$

where r_1 and r_2 are the principal radii of curvature. Since now

$$(12)\qquad h - k = \frac{1}{2}\,\frac{\partial^2 \log E/G}{\partial u\,\partial v}, \qquad h' - k' = \frac{1}{2}\,\frac{\partial^2 \log \mathfrak{E}/\mathfrak{G}}{\partial u\,\partial v},$$

we find, from the second equation of (11), that

$$(13)\qquad \frac{\partial^2 \log r_1/r_2}{\partial u\,\partial v} = (h-k) - (h'-k').$$

But relations (6) become

$$(14)\qquad \frac{\partial \log E}{\partial v} = \frac{r_2}{r_1}\,\frac{\partial \log \mathfrak{E}}{\partial v}, \qquad \frac{\partial \log G}{\partial u} = \frac{r_1}{r_2}\,\frac{\partial \log \mathfrak{G}}{\partial u},$$

and so (13) can be written in the form

$$(15)\qquad 2\,\frac{\partial^2 \log r_1/r_2}{\partial u\,\partial v} = \frac{\partial}{\partial u}\left[\left(\frac{r_2}{r_1} - 1\right)\frac{\partial \log \mathfrak{E}}{\partial v}\right] - \frac{\partial}{\partial v}\left[\left(\frac{r_1}{r_2} - 1\right)\frac{\partial \log \mathfrak{G}}{\partial u}\right],$$

which is the equation to which (5) reduces in this case.

THEOREM 7. *An orthogonal system of curves on the sphere represents the lines of curvature on a surface for which the ratio r_1/r_2 of the principal radii of curvature is prescribed, if and only if $\mathfrak{E}$, $\mathfrak{G}$, r_1/r_2 satisfy* (15). *The surface is then determined to within its homothetics.*

Consider now the four properties: (a) the ratio of the principal radii of curvature of the form $U(u)/V(v)$, the lines of curvature being para-

metric; (b) the lines of curvature an isothermal-orthogonal system; (c) the lines of curvature an isothermal-conjugate system; (d) the spherical representation of the lines of curvature an isothermal system. From (11) follows immediately the theorem:

THEOREM 8. *If a surface has any two of the properties enumerated, it has the other two also.**

If $r_1/r_2 = -1$, we find, from (15), that $\mathfrak{E}/\mathfrak{G}$ is a function of the form $U(u)/V(v)$.

THEOREM 9. *The lines of curvature of a minimal surface are isothermal, both as an orthogonal and as a conjugate system, and are represented on the sphere by an isothermal system. Conversely, an isothermal (orthogonal) system on the sphere represents the lines of curvature on a minimal surface unique to within its homothetics.*

From (15), in conjunction with (11), we obtain the following results.

THEOREM 10. *An orthogonal system of curves on the sphere represents the lines of curvature on a surface of Bour for which the ratio r_1/r_2 is prescribed, if and only if $\frac{\mathfrak{E}}{\mathfrak{G}}\left(\frac{r_1}{r_2}\right)^2$ is of the form $U(u)/V(v)$ and $\mathfrak{E}$, $\mathfrak{G}$, r_1/r_2 satisfy*

$$(16)\qquad \frac{\partial}{\partial u}\left(\frac{r_2}{r_1}\frac{\partial \log \mathfrak{E}}{\partial v}\right) = \frac{\partial}{\partial v}\left(\frac{r_1}{r_2}\frac{\partial \log \mathfrak{G}}{\partial u}\right).$$

If, however, the lines of curvature are to form an isothermal-conjugate system, it is $\frac{\mathfrak{E}}{\mathfrak{G}}\frac{r_1}{r_2}$ which is to be of the form $U(u)/V(v)$; in this case $\mathfrak{E}$, $\mathfrak{G}$, r_1/r_2 must satisfy

$$\frac{\partial}{\partial u}\left[\left(\frac{r_2}{r_1}+1\right)\frac{\partial \log \mathfrak{E}}{\partial v}\right] = \frac{\partial}{\partial v}\left[\left(\frac{r_1}{r_2}+1\right)\frac{\partial \log \mathfrak{G}}{\partial u}\right].$$

Finally, if the lines of curvature are to be isothermal, both as an orthogonal and a conjugate system, both $\mathfrak{E}/\mathfrak{G}$ and r_1/r_2 must be of the form $U(u)/V(v)$ and $\mathfrak{E}$, $\mathfrak{G}$, r_1/r_2 must satisfy (16).

THEOREM 11. *A necessary and sufficient condition that an isothermal (orthogonal) system of curves on the sphere represent an isothermal-orthogonal or*

---

* Eisenhart has shown that if a surface has any two of the last *three* properties, it has the third also, "Isothermal-conjugate lines on surfaces", Amer. Journ. Math., vol. 25 (1903), pp. 213—248, in particular, p. 228. From Eisenhart's result can be deduced the first part of Theorem 9, since it is well known that the lines of curvature of a minimal surface and their spherical representation form isothermal-orthogonal systems.

isothermal-conjugate system of lines of curvature on a surface for which the ratio r_1/r_2 is prescribed is that r_1/r_2 be of the form $U(u)/V(v)$ and $\mathfrak{E}$, $\mathfrak{G}$, r_1/r_2 satisfy (16).

Equation (13), which is the equivalent in this case of the general equation (7), embodies the following result.

THEOREM 12. *When the lines of curvature on a surface are parametric, the ratio of the principal radii of curvature is of the form $U(u)/V(v)$ if and only if the difference of the point invariants of the lines of curvatare equals the difference of their plane invariants.*

7. Asymptotic lines. If a surface S of negative curvature is referred to one family of asymptotic lines as the u-curves and to any other family of real curves as the v-curves, the Codazzi equations become

$$\frac{\partial f}{\partial u} = \left[\left\{{11 \atop 1}\right\}' - \left\{{12 \atop 2}\right\}'\right] f + \left\{{11 \atop 2}\right\}' g.$$

$$\frac{\partial f}{\partial v} - \frac{\partial g}{\partial u} = \left[\left\{{22 \atop 2}\right\}' - \left\{{12 \atop 1}\right\}'\right] f - \left\{{12 \atop 2}\right\}' g.$$

We divide each equation by f, replace the second by the equation obtained by adding the first multiplied by g/f to the second, and simplify by means of the identities,

$$\frac{\partial \log \mathfrak{H}}{\partial v} = \left\{{22 \atop 2}\right\}' + \left\{{12 \atop 1}\right\}', \qquad \frac{\partial \log \mathfrak{H}}{\partial u} = \left\{{11 \atop 1}\right\}' + \left\{{12 \atop 2}\right\}'.$$

The result is

$$-\frac{\partial \log f/\mathfrak{H}}{\partial u} = 2\left\{{12 \atop 2}\right\}' - \left\{{11 \atop 2}\right\}' \frac{g}{f},$$

(17)

$$\frac{\partial \log f/\mathfrak{H}}{\partial v} = \frac{\partial}{\partial u}\frac{g}{f} - 2\left\{{12 \atop 1}\right\}' + \left[\left\{{11 \atop 1}\right\}' - 2\left\{{12 \atop 2}\right\}'\right]\frac{g}{f} + \left\{{11 \atop 2}\right\}'\left(\frac{g}{f}\right)^2.$$

Hence

$$(18) \quad \frac{\partial^2}{\partial u^2}\frac{g}{f} - \frac{\partial}{\partial u}\left\{2\left\{{12 \atop 1}\right\}' - \left[\left\{{11 \atop 1}\right\}' - 2\left\{{12 \atop 2}\right\}'\right]\frac{g}{f} - \left\{{11 \atop 2}\right\}'\left(\frac{g}{f}\right)^2\right\}$$

$$+ \frac{\partial}{\partial v}\left[2\left\{{12 \atop 2}\right\}' - \left\{{11 \atop 2}\right\}'\frac{g}{f}\right] = 0.$$

Since $e = 0$, the second family of asymptotic lines is defined by the equation

$$du + M\,dv = 0,$$

where

$$M = \frac{g}{2f}. \tag{19}$$

Equation (18), written in terms of M, is

$$\frac{\partial^2 M}{\partial u^2} - \frac{\partial}{\partial u}\left\{\begin{Bmatrix}12\\1\end{Bmatrix}' - \left[\begin{Bmatrix}11\\1\end{Bmatrix}' - 2\begin{Bmatrix}12\\2\end{Bmatrix}'\right]M - 2\begin{Bmatrix}11\\2\end{Bmatrix}' M^2\right\} \tag{20}$$

$$+ \frac{\partial}{\partial v}\left[\begin{Bmatrix}12\\2\end{Bmatrix}' - \begin{Bmatrix}11\\2\end{Bmatrix}' M\right] = 0.$$

Conversely, if there is given on the Gauss sphere a system of curves, parametric, and a point function M, such that $\mathfrak{E}$, $\mathfrak{F}$, $\mathfrak{G}$, M satisfy (20), equations (17) in f, where g/f has been replaced, by $2M$, are integrable by a quadrature and determine f except for a multiplicative constant, k; $g = 2fM$ is then unique except for the same multiplier, k. Consequently, the surface S is determined to within its homothetics. Its point coördinates, when those of the sphere are known, can be found by quadratures from the equations

$$\frac{\partial x}{\partial u} = \frac{f}{\mathfrak{H}^2}\left(\mathfrak{F}\frac{\partial \zeta}{\partial u} - \mathfrak{E}\frac{\partial \zeta}{\partial v}\right), \quad \frac{\partial x}{\partial v} = \frac{1}{\mathfrak{H}^2}\left[(\mathfrak{F}g - \mathfrak{G}f)\frac{\partial \zeta}{\partial u} + (\mathfrak{F}f - \mathfrak{E}g)\frac{\partial \zeta}{\partial v}\right],$$

and are determined except for the multiplier k and additive constants.

THEOREM 13. *The u-curves of a real parametric system on the sphere represent one family of asymptotic lines on a surface for which the differential equation $du + M\,dv = 0$ of the second family of asymptotic lines is prescribed if and only if $\mathfrak{E}$, $\mathfrak{F}$, $\mathfrak{G}$, M satisfy* (20). *The surface is then determined to within its homothetics and its point coördinates can be found by quadratures, when those of the sphere are known.*

If $M = 0$, that is, if the given system on the sphere is to represent both families of asymptotic lines, condition (20) reduces to

$$\frac{\partial}{\partial u}\begin{Bmatrix}12\\1\end{Bmatrix}' = \frac{\partial}{\partial v}\begin{Bmatrix}12\\2\end{Bmatrix}'.$$

We thus obtain Dini's Theorem* as a special case.

* Cf. Introduction.

Since, when $e = 0$,

$$\left\{\begin{matrix}11\\2\end{matrix}\right\}' = -\left\{\begin{matrix}11\\2\end{matrix}\right\},$$

and the vanishing of the latter symbol is a necessary and sufficient condition that the u-curves be straight, we conclude the following:

THEOREM 14. *The u-curves of a real parametric system on the sphere represent the rulings of a ruled surface for which the differential equation $du + M\,dv = 0$ of the second family of asymptotic lines is prescribed, if and only if* $\left\{\begin{matrix}11\\2\end{matrix}\right\}' = 0$ *and* $\mathfrak{E}, \mathfrak{F}, \mathfrak{G}, M$ *satisfy the equation*

$$\frac{\partial^2 M}{\partial u^2} - \frac{\partial}{\partial u}\left\{\left\{\begin{matrix}12\\1\end{matrix}\right\}' - \left[\left\{\begin{matrix}11\\1\end{matrix}\right\}' - 2\left\{\begin{matrix}12\\2\end{matrix}\right\}'\right] M\right\} + \frac{\partial}{\partial v}\left\{\begin{matrix}12\\2\end{matrix}\right\}' = 0. \tag{21}$$

If the given surface is minimal, $M = \mathfrak{F}/\mathfrak{E}$. Consequently, the u-curves of a system on the sphere represent one family of asymptotic lines on a minimal surface if and only if, when M in (20) is replaced by $\mathfrak{F}/\mathfrak{E}$, the equation is satisfied.

HARVARD UNIVERSITY,
CAMBRIDGE, MASS.,
June, 1921.

ON A SHORT METHOD OF LEAST SQUARES.

By Burton H. Camp.

1. **Introduction.** In one of the "Scientific Papers of the Bureau of Standards"* it is shown that, when the independent variables of a set of observational equations are equi-spaced, certain simplifications are available which greatly shorten the least square solution. These simplifications are of the same character as those which have been used by statisticians in computing the coefficients of regression,† and the trend in a time series.‡ They may be described in a particular case as follows:

Consider the type equation:

$$Z = A + BX + CY, \tag{1}$$

where X and Y are independent variables, giving n equations of condition to determine A, B and C. To obtain the first normal equation, add the equations of condition. To obtain the second, arrange them in order of ascending X's, multiply the successive equations through by the successive terms of any ascending, equi-spaced sequence of integers, and then add the results. To obtain the third normal equation, arrange in order of ascending Y's, multiply by a sequence of integers, as before, and add. It is sometimes supposed necessary that the sequence of integers shall be special in form, but this is in fact immaterial, all ascending, equi-spaced sequences leading to the same results. It will usually be convenient, however, to choose the sequence:

$$-n+1,\ -n+3,\ \ldots,\ -1,\ +1,\ \ldots,\ +n-1, \tag{2}$$

when n is even; and the sequence:

$$-(n-1)/2,\ -(n-1)/2+1,\ \ldots,\ 0,\ \ldots,\ (n-1)/2, \tag{3}$$

* No. 388, Adjustment of parabolic and linear curves to observations taken at equal intervals of the independent variable, by Harry M. Roeser, July, 1920, Department of Commerce, U. S. A.

† E. g., Yule, Theory of statistics, Griffin and Co., Ltd., 1911, pp. 181 ff.

‡ E. g., Persons, The Review of economic statistics, Preliminary Volume I, 1919, p. 13, foot-note 1.

when n is odd. These sequences will be assumed in the calculations to follow and in the resulting expressions (10) for the errors. It will be shown that, if the X's and the Y's are not equi-spaced but are approximately so, this short method may be used with negligible error, and expressions for the error will be found. It often happens that only approximate equi-spacing is possible. One may have only partial control of the independent variables, X, Y, and it would be very difficult or impossible to choose them exactly equi-spaced, as, for example, when selecting a star list for time observations with a transit. Sometimes, when one does have apparent control of these variables and has set his instrument so that they are equi-spaced, the spacing is disturbed after allowance is made for instrumental errors; the same effect is produced if the observations have slightly varying weights. It is, therefore, desirable that an expression for the error in using the short method should be obtained, and it should be in such a form that it may be computed quickly, after the short determination is made. Further, it should be possible to estimate roughly the size of the error to be expected, before *any* computation has been made.

If the order of the X's is the same as the order of the Y's the method as outlined above fails. An important case where this happens is when $Y=X^2$ in the type equation. It is feasible to make special arrangements in such cases, but they will not be discussed here. The results in the special cases, A, B, or C equal to zero, will be given, however. The computer may use the summary in § 3 below without reading the theory in § 2.

2. **Theory.** Let i be the representative integer of the sequence (2), or (3), while the second normal equation is being formed, and j the representative integer for the third normal equation. Let X_0 denote the middle X, when n is odd, the point half way between the two middle X's, when n is even. Let h denote the "normal" space interval of the X's, this to mean that value which will make the sum of the squares of the errors (e) made in writing, $X-X_0=hi$, a minimum. For the Y's, let Y_0, k, and d be the analogous letters. Let r denote the residual, $Z-(A+BX+CY)$. Thus,

$$(4)\qquad h=\frac{\sum i(X-X_0)}{\sum i^2},\qquad k=\frac{\sum j(Y-Y_0)}{\sum j^2}.$$

The true normal equations, resulting from an exact application of the least square method may be written

$$(5)\quad \sum r=0,\quad \sum(hi+e+X_0)\,r=0,\quad \sum(kj+d+Y_0)\,r=0,$$

and, by virtue of the first of these equations, the $\sum X_0\,r$ and $\sum Y_0\,r$ of the second and third drop out. Make the transformations

(6) $X = x + \bar{X},\ Y = y + \bar{Y},\ Z = z + \bar{Z},$ where $\sum x = \sum y = \sum z = 0.$

Equations (5) become, after reduction,

$$
\begin{aligned}
\bar{Z} &= A + B\bar{X} + C\bar{Z}, \\
(7) \qquad \sum hiz &= B\sum hix + C\sum hiy - \sum er, \\
\sum kjz &= B\sum kjx + C\sum kjy - \sum dr.
\end{aligned}
$$

Solving by determinants the last two of (7), one gets the true values of B and C; putting $e = d = 0$ in the results gives the values of B and C that would be obtained by the approximate method under consideration. The differences will be denoted by dB and dC, and are

$$
(8) \qquad
\begin{aligned}
dB &= \frac{1}{\Delta}\begin{vmatrix} \sum er & h\sum iy \\ \sum dr & k\sum jy \end{vmatrix}, \qquad dC = \frac{1}{\Delta}\begin{vmatrix} h\sum ix & \sum er, \\ k\sum jx & \sum dr, \end{vmatrix}, \\
\Delta &= hk\begin{vmatrix} \sum ix & \sum iy \\ \sum jx & \sum jy \end{vmatrix}.
\end{aligned}
$$

It is now necessary to find approximate values of these expressions in terms of quantities which will be found in the course of the short method computation, or of other known quantities.

Let s_A, s_B, s_C refer to the mean square deviations*, called by the statistician standard deviations, of A, B, C, respectively, due to errors of sampling. They may be found approximately by using the values of B and C obtained by the approximate method. Thus, by the ordinary theory of errors,

$$
\begin{aligned}
s_B^2 &= \frac{h^2k^2}{\Delta^2}\left[(s_r^2\sum i^2)\ (\sum jy)^2 + (s_r^2\sum j^2)\ (\sum iy)^2\right] \\
&= \frac{h^2k^2s_r^2}{\Delta^2}\left[(\sum jy)^2 + (\sum iy)^2\right]\sum i^2,
\end{aligned}
$$

where $s_r = \sqrt{\frac{1}{n}\sum r^2}$, and this also equals the mean square deviation of the z's, about the mean plane. Solving this equation for Δ^2 and substituting in (8), one gets

* One might as well use "probable" deviations, or probable errors, by multiplying the standard deviations by the appropriate constants.

$$(9) \qquad dB = \frac{\begin{vmatrix} \sum er & h\sum iy \\ \sum dr & k\sum jy \end{vmatrix}}{hks_r\sqrt{(\sum jy)^2 + (\sum iy)^2}} \cdot \frac{s_B}{\sqrt{\sum i^2}}.$$

Now e and r are uncorrelated, and so the most probable value of their coefficient of correlation, $\sum er / s_e s_r n$, where $s_e = \sqrt{\sum e^2/n}$, is zero, and it is almost certain* that this quantity does not depart from zero by more than three times its own standard deviation; it is probable that this departure is, in fact, less than the standard deviation. Thus, almost certainly,*

$$\left|\frac{\sum er}{n\, s_e\, s_r}\right| < \frac{3}{\sqrt{n}},$$

and therefore $\left|\sum er\right| < 3\, s_r\, s_e \sqrt{n}$; and similarly $\left|\sum dr\right| < 3\, s_r\, s_d \sqrt{n}$. On substitution in (9), it is found that, almost certainly,†

$$|dB| < 3s_B\sqrt{\frac{n}{\sum i^2}} \cdot \frac{s_e\left|\sum jy\right| k + s_d\left|\sum iy\right| h}{hk\sqrt{(\sum jy)^2 + (\sum iy)^2}}.$$

An approximation to this expression may be made by putting $y = Y - Y_0$, since $Y - Y_0$ is small. The result may then be simplified by using the fact that $\sum i = \sum j = 0$, and inserting the values of h and k from (4). Finally, then,

$$(10)\dagger \qquad \begin{aligned} |dB| &< 3s_B\sqrt{n\sum i^2}\, \frac{\left|\dfrac{\sum jY}{\sum iX}\right| s_e + \left|\dfrac{\sum iY}{\sum jY}\right| s_d}{\sqrt{(\sum jY)^2 + (\sum iY)^2}}, \\ |dC| &< 3s_B\sqrt{n\sum i^2}\, \frac{\left|\dfrac{\sum iX}{\sum jY}\right| s_e + \left|\dfrac{\sum jX}{\sum iX}\right| s_d}{\sqrt{(\sum iX)^2 + (\sum jX)^2}}, \\ dA &= -\frac{1}{n}\left(dB\sum X + dC\sum Y\right). \end{aligned}$$

---

* On the assumption of a Gaussian distribution, the probability that the coefficient of correlation will exceed its standard deviation is 0.32, and the probability that it will exceed three times its standard deviation is only 0.0027.

† See preceding foot-note. In the examples following the theory, limits obtained from (10) are described as "nearly certain". Limits obtained from (10) without the factor 3 are described as "probable".

These formulae hold good only if the i and j sequences are chosen as in (2) and (3). All the sums involved occur in the normal equations as determined by the short method. After these sums have been found, a slide rule computation suffices for dB/s_B, and dC/s_C. It is these ratios, and not the absolute values of the errors which are important in determining the question whether the short method is justified. For example, if $\lambda = dB/s_B$ is as small as 0.01, then there is a probability of 0.996 that dB, the error introduced by the short method, is numerically less than the necessary error in B, due to the accidental errors of observation. The corresponding probabilities for other λ's are given by a Gaussian table, as follows:

λ	prob.	λ	prob.	λ	prob.
.005	.996	.10	.92	.4	.69
.01	.992	.20	.84	.5	.62
.05	.957	.30	.76	1.0	.32

Similar remarks apply to C. As will become obvious from the examples to follow, it is necessary to have only an approximate estimate of the sizes of s_e and s_d. These are sometimes obvious from the nature of the problem (e. g., Ex. 2 below). In other cases they must be computed. The labor of doing this is slight. Consider, for example, s_e. The values of $\sum iX$ and $\sum i^2$ from which h is found have already been obtained in the short solution. Then $e_i = X_i - X_0 - hi$, $s_e^2 = \sum e^2/n$. It follows from the last equation of (10) that dA may be made zero, *whether the X's and Y's are nearly equi-spaced or not*, provided they are so chosen that $\sum X = \sum Y = 0$. This is a consideration worth keeping in mind if A is the quantity most desired from the observations. It follows also that, if $\sum X/n$ and $\sum Y/n$ are small, the approximate solution for A may be used, instead of the long solution, with small error. It seems to the author that this method could be used to advantage in getting the time from transit observations. Tables already in use in connection with other methods make it easy to select the stars properly (Cf. Ex. 3 below).

The expressions just found are available after the approximate normal equations have been found. It is now desirable to shorten them by further approximations so that they may be made to yield good estimates of the errors, in advance of *any* computation. It will be necessary to assume now that the order of the X's is uncorrelated with the order of the Y's. This is frequently the case, but there are important instances in which it is not, and in such instances the subsequent approximations must not be used. Approximately,

(11) $\sum i(Y - Y_0) = \sum j(X - X_0) = 0, \; X_i - X_0 = ih, \; Y_j - Y_0 = jk.$

Omitting the coefficients 3 in (10), and substituting (11), one obtains, as a provisional value for dB,

(12) $$dB = \frac{s_B s_e}{h} \sqrt{\frac{n}{\sum i^2}}.$$

If n is odd, $\sum i^2 = n(n-1)(n+1)/12$. If n is even, $\sum i^2 = n(n-1)(n+1)/6$ but in the latter case h and k are only half what they are in the former case, to that, if H and K be approximately the average intervals between the X's, and between the Y's, respectively, then in all cases:

(13)* $$\frac{|dB|}{s_B} = \frac{s_e}{H} \cdot \frac{2\sqrt{3}}{\sqrt{(n+1)(n-1)}}, \qquad \frac{|dC|}{s_C} = \frac{s_d}{K} \cdot \frac{2\sqrt{3}}{\sqrt{(n+1)(n-1)}}.$$

The quantities s_e/H, s_d/K, measure the divergences from equi-spacing, and it is natural than they should appear as factors in the expressions for the errors. In most problems where any real attempt is made at equi-spacing it is possible to make s_e/H and s_d/K as small as 1/3. Indeed, ten times this smallness is to be expected frequently. In very bad cases, as in Example 3, where the conditions of the problem are such as to interfere with equi-spacing, these quantities may be as large as 1.5. It follows from (13) that dB/s_B and dC/s_C approach zero with $1/n$. One must compute dA from the formula of (10).

3. **Summary.** Let dA, dB, dC be the errors made in A, B, C if the short method of § 1 is used in connection with (1) instead of the exact least square solution. To find dB/s_B and dC/s_C, provisionally, substitute in (13), s_B and s_C being the mean square deviations† of B and C due to the fluctuations of the observations themselves. The values of s_e/H and s_d/K may be taken as stated under (13). To find dA/s_A, then use the last equation of (10). Formulae (13) are independent of the particular type of ascending equi-spaced sequence of integers used as multipliers, but formulae (10) are not.

To find closer approximations to these relative errors, after the short computation has been made, use equations (10), noting what is said in the remainder of the paragraph containing (10). $\sum i^2$ is best found from a short table which the computer may as well make for himself, although several such tables are published for large values of n‡.

---

* Probable values. Insert the factors 3 to obtain more certain limits. See preceding foot-note.

† Also called "standard deviations".

‡ E. g., "Tables for statisticians and biometricians", by Karl Pearson, Cambridge University Press, 1914.

For the special case $Z = BX + CY$, the same formulae may be used if, and only if, X_0 and Y_0 are numerically small, relative to those values of X and Y which are distant from X_0 and Y_0. Otherwise, it is not recommended that the short method be used.

For the special case $Z = A + BX$, formulae (10) and (13) take on the simple forms:

$$(10a)^* \qquad \frac{|dB|}{s_B} = \sqrt{n\sum i^2} \cdot \frac{s_e}{\left|\sum iX\right|}, \qquad dA = -dB\frac{\sum X}{n},$$

$$(13a)^* \qquad \frac{|dB|}{s_B} = \frac{s_e}{H} \cdot \frac{2\sqrt{3}}{\sqrt{(n+1)(n-1)}},$$

and the various remarks made in connection with (10) and (13) apply of course to these special forms.

4. **Examples.** *Example 1.* For the first illustration a rather bad case will be chosen, a group of five stars composing a time set, with only the crudest attempt at equi-spacing. They are chosen from the unweighted set in Campbell's "Elements of Practical Astronomy", page 150, and in the following observational equations, the letters A, B, and C, correspond with Campbell's x, c, and a, respectively. The equations are

$$\begin{aligned}
-1.66 &= A - 3.813\,B + 3.395\,C \\
-0.13 &= A + 1.007\,B + 0.587\,C \\
+0.98 &= A + 2.915\,B - 1.353\,C \\
-0.02 &= A - 1.714\,B - 0.357\,C \\
+0.47 &= A - 4.898\,B - 2.875\,C
\end{aligned}$$

The results, as obtained by the exact method of least squares and also by the approximate method are:

	True Values	Approximate Values	Differences, called "errors" above	Standard Deviations
A	.0488	.0512	.0024	.024
B	.1266	.1289	.0023	.008
C	—.3658	—.3689	.0030	.011

* Probable values. See foot-notes to (10) and (13).

The next table gives the provisional and final estimates of the maximum values of the errors, as derived from the formulae of this paper, also their true values. In using equation (13) it was assumed that $s_e/H = s_d/K = 1/3$. It turned out that these numbers were $0._{27}$ and $0._{36}$, respectively. From the table in § 2, it follows that, assuming a normal distribution of error, the fluctuation of A due to the observations themselves, would, in 92 cases out of 100, exceed the difference ($.{\scriptstyle 0024}$) introduced by the short method of computation, and the corresponding numbers for B and C are 77 and 78 cases. It would seem to the author that, in examples like this, the short method is usually justifiable, but that it might not be sometimes. It is, however, not the object to present an argument for the short method, but to present a criterion by which the error to be committed by it may be estimated.

	Limits* by (13)		Limits by (10)		True Values
	Probable	Nearly Certain	Probable	Nearly Certain	
$\lvert dA/s_A \rvert$	—	—	—	—	.104
$\lvert dB/s_B \rvert$	.24	.72	.20	.60	.287
$\lvert dC/s_C \rvert$	.24	.72	.26	.75	.282

Example 2. A very good case will now be considered, a set of observations used by Roeser† with the equi-spacing disturbed by the arbitrary addition of small quantities. These small quantities were chosen at random from a table of numbers subject to the limitation that they should lie between $.{\scriptstyle 05}$ and $-.{\scriptstyle 05}$. The equations and tables following have interpretations similar to those of Example 1. It is obvious from a mere glance at the equations that s_e/H is not greater than about $.{\scriptstyle 004}$, and this was the value used in the provisional estimate. The true value is $0.{\scriptstyle 0028}$. If this had been used instead of $0.{\scriptstyle 004}$, the provisional estimate for dB/s_B would have differed from the true value by only about 4 per cent., thus illustrating the closeness of the inequality. It is not desirable, however, in practice to compute s_e/H in advance, and accordingly the rough estimate ($0.{\scriptstyle 004}$) is used. With this rough approximation, formula (13) shows that the value of dB/s_B will be so small as to make the difference between the two methods entirely negligible. This is exhibited in the table below.

* The first column is obtained from (13). The second column is three times the first. See also foot-note to (10).

† Loc. cit., p. 371. See also Weld, Theory of errors and least squares, 1916, p. 76.

Conditional Equations

$$
\begin{aligned}
0.116 &= A + 10.05\,B, & \qquad 0.595 &= A + 60.01\,B,\\
0.205 &= A + 19.95\,B, & 0.675 &= A + 69.98\,B,\\
0.295 &= A + 29.99\,B, & 0.760 &= A + 79.98\,B,\\
0.388 &= A + 40.01\,B. & 0.850 &= A + 90.04\,B,\\
0.503 &= A + 49.97\,B, & 0.926 &= A + 100.01\,B.
\end{aligned}
$$

	True	Approximate	Difference	Standard Deviation
A	.028478	.028425	.000053	.00315
B	.0091417	.0091418	.0000001	.00011

	Limits by (13a)		Limits by (10a)		True
	Probable	Nearly Certain	Probable	Nearly Certain	
dA/s_A	—	—	—	—	.00168
dB/s_B	.00139	.00417	.00099	.00297	.00096

A is not so well determined as B. This is because the origin is far from the middle X.

Example 3. Finally, a very bad case will be considered, viz., Campbell's eleven weighted equations* out of which the data of Example 1 were selected. Here the method of choice of stars was such as deliberately to prevent equi-spacing, as a rough plotting of U, X, and Y would show. The typical equation is

$$Z = AU + BX + CY,$$

a type not considered in the theory, but formulae (13) are valid. In such bad cases as this, the divergence from equi-spacing is so great that the ratios like s_e/H, which measure it, cannot be assumed small, and so it is advisable to estimate them by a graph. If, for example, X_i be plotted as a function of i, it will be seen that s_e is about unity and H about 2/3, making s_e/H about 1.5. The corresponding fractions for U and Y may be found in a similar manner to be about 1.2 and 1.0, respectively.

* Loc. cit., p. 155.

	True	Approximate	Difference	Standard Deviation		Limits by (13)		True
						Probable	Nearly Certain	
A	.053	.053	.000	.017	dA/s_A	.39	1.2	.01
B	.115	.122	.007	.012	dB/s_B	.47	1.3	.58
C	−.383	−.386	.003	.020	dC/s_C	.32	1.0	15

Thus, even in this case, the provisional formula (13) is adequate to show that the errors are probably as small as half the standard deviations of the quantities measured. The very small actual error in the time correction, A, is due in part to the fact that the means of the X's and Y's are small. This example, taken together with Example 1, also serves to illustrate how the disturbance due to bad spacing can be masked by an increase in the number of equations of condition. This result is to be expected from the form of (13), as was noted in the sentence preceding the last of § 2. The value of the short method in saving labor increases greatly with the number of equations of condition, and it is therefore particularly comforting to know that at the same time the error in using it diminishes.

5. **Conclusion.** The short method works sufficiently well to justify its use, certainly in cases as good as Example 2, probably in cases like Example 1, and perhaps (depending on the accuracy desired) in cases as bad as Example 3. In any case the goodness of the method may, with sufficient accuracy, be judged in advance by the criteria (13) or (13a). When the short method is used, the gain in time is very great, especially if one employs a computing machine. The author used an electrically driven listing machine, but an ordinary adding machine would have done almost as well. All the twelve coefficients of the normal equations in Example 3 for the eleven equations of condition were found in this way in less than twenty minutes. In the short method, all the multiplications necessary for the coefficients like $\sum iX$ are by small integers. Thus, with an adding machine, the products, iX, may be found and added at the same time.

Wesleyan University,
May, 1922.

ON THE CONVERGENCE OF THE STURM-LIOUVILLE SERIES.*

By J. L. Walsh.

1. We shall consider in this paper the differential equation

$$u''(x)+[\varrho^2-g(x)]\,u(x) = 0, \qquad 0 \leqq x \leqq 1, \tag{1}$$

where ϱ is a parameter, in connection with the boundary conditions

$$u(0) = 0, \qquad u(1) = 0. \tag{2}$$

For certain values of ϱ, the so-called characteristic values, equation (1) has solutions (normal solutions or characteristic functions) which satisfy (2), and in mathematical physics there arises the problem of the development of arbitrary functions in series in terms of these normal solutions. These series are a special, but important and typical, case of the more general Sturm-Liouville series which are the developments of arbitrary functions in terms of normal solutions of equations similar to (1) and which satisfy homogeneous boundary conditions similar to (2).

Under certain restrictions on $g(x)$, we shall prove that, for any function on the interval $0 \leqq x \leqq 1$ integrable in the sense of Lebesgue and with an integrable square, the series which is the formal expansion in terms of the normal functions which correspond to (1), (2) has essentially the same convergence properties as the series which is the formal expansion in terms of the normal functions which correspond to (1), (2) when $g(x) \equiv 0$. This last set of functions is, except for a constant factor, the set $\{\sin k\pi x\}$, and the convergence properties of the expansions of arbitrary functions in terms of this set are well known.

2. The results, but not the methods, of the present paper are closely connected with the work of Haar† although he considers the boundary conditions

$$u'(0)-hu(0) = u'(1)+Hu(1) = 0 \tag{3}$$

---

* Presented to the American Mathematical Society, December, 1920. It is largely due to Dr. T. H. Gronwall that the result here published has its present comparatively simple form. Thus Dr. Gronwall eliminated from the hypothesis of the principal theorem an unnecessary assumption and indicated to the writer more than a corresponding simplification of the use of the asymptotic expressions.

† Math. Annalen, vol. 69 (1910), pp. 331–371; vol. 71 (1912), pp. 38–53.

instead of conditions (2) in connection with equation (1). Thus Haar proves that the Sturm-Liouville series of an integrable function is convergent, divergent, or summable at a point according as its Fourier cosine series is convergent, divergent, or summable at that point; the present paper deals merely with functions integrable and with an integrable square. Haar does not, however, bring out clearly the identity of the properties of uniform convergence for the two developments,* nor is it obvious how his methods can be extended to include absolute convergence; the present paper deals with both uniform and absolute convergence. Haar makes use of the fact that an analytic function can be approximated as closely as desired by a linear combination of normal functions; the present paper proves at a single step the possibility of expansion of all functions (integrable and with an integrable square) which can be developed into a Fourier sine series, and the identity of the convergence properties of the two developments.

The similarity of uniform convergence for the two expansions, in the precise manner which appears later, enables us to state that Gibbs' phenomenon occurs for our Sturm-Liouville series in precisely the same manner as for the Fourier sine series. Gibbs' phenomenon for the Sturm-Liouville series seems first to have been pointed out by Weyl.†

3. We shall first prove a general theorem concerning expansions of arbitrary functions and then apply that theorem to the system (1), (2).

A set of functions $\{u_n(x)\}$ $(n = 1, 2, 3, \ldots)$ continuous on the interval $0 \leq x < 1$ is said to be *normal* on that interval if and only if

$$\int_0^1 u_n^2(x)\,dx = 1, \qquad (n = 1, 2, 3, \ldots),$$

and is said to be *orthogonal* if and only if

$$\int_0^1 u_i(x)\,u_j(x)\,dx = 0, \qquad (i \neq j).$$

We shall consider, under certain restrictions, two normal orthogonal sets of functions $\{u_n(x)\}$ and $\{\bar{u}_n(x)\}$ and shall suppose that the $\{\bar{u}_n(x)\}$ can be

* For uniform convergence and the developments of *continuous* functions, see the comment on Haar's work by Bôcher, Proceedings of the Fifth International Congress of Mathematicians (Cambridge), vol. I, pp. 191, 192.

Compare also the reference given in the next following footnote.

† For equation (1) in connection with the boundary conditions $u'(0) = u'(1) = 0$, see Rend. Circ. Mat. Palermo, vol. 29 (1910), pp. 321–323.

expanded in terms of the $\{u_n(x)\}$. We invert this system of equations and thereby express the $\{u_n(x)\}$ in terms of the $\{\bar{u}_n(x)\}$. To prove that the convergence properties of the formal expansion of an arbitrary function $f(x)$ in terms of the $\{\bar{u}_n(x)\}$ are essentially the same as the convergence properties of the expansion in terms of the $\{u_n(x)\}$ (these properties are assumed to be known), in the former expansion we merely substitute the expansion of the $\{\bar{u}_n(x)\}$ in terms of the $\{u_n(x)\}$. Rearrangement of the terms then gives us precisely the expansion of $f(x)$ in terms of the $\{u_n(x)\}$. The following exposition will not seem unnatural if this general method is kept in mind.

4. To prepare for the inversion of the system of equations indicated, we now prove

LEMMA I. *If the set of real numbers* $\{c_{nk}\}$, $n, k = 1, 2, 3, \ldots,$ *is such that*

$$c_{nk} + c_{kn} + \sum_{\nu=1}^{\infty} c_{n\nu}\, c_{k\nu} = 0, \tag{4}$$

and if the series

$$\sum_{n=1}^{\infty} \left| \sum_{k=1}^{\infty} |c_{nk}| \right|^2 \tag{5}$$

converges, then we have also

$$c_{kn} + c_{nk} + \sum_{\nu=1}^{\infty} c_{\nu n}\, c_{\nu k} = 0. \tag{6}$$

We shall have frequent occasion to use the Lagrange inequality which holds for any two sets of real numbers $\{a_\nu\}$ and $\{b_\nu\}$:

$$\left(\sum_{\nu=1}^{m} |a_\nu\, b_\nu|\right)^2 \leq \sum_{\nu=1}^{m} a_\nu^2 \sum_{\nu=1}^{m} b_\nu^2,$$

whence the well known and frequently used fact that, when $\sum_\nu a_\nu^2$ and $\sum_\nu b_\nu^2$ both converge, then also $\sum_\nu |a_\nu\, b_\nu|$ converges and

$$\left(\sum_\nu |a_\nu\, b_\nu|\right)^2 \leq \sum_\nu a_\nu^2 \sum_\nu b_\nu^2. \tag{7}$$

Here and below, unless otherwise stated, summation subscripts run from 1 to ∞.

From the convergence of (5), then, follows immediately the convergence of the triple series

$$(8) \qquad \sum_{\mu\nu n} |c_{n\mu} c_{n\nu}|.$$

All terms of the series

$$(9) \qquad \sum_{nk} c_{nk}^2,$$

$$(10) \qquad \sum_{\nu} c_{n\nu}^2,$$

are contained in (8), so these two series converge. The absolute convergence of the series contained in (4) follows by application of (7). The absolute convergence of the series contained in (6) follows from the convergence of (8).

Denote by ϵ_{nk} the left-hand member of (6), so that we have

$$(11) \quad \epsilon_{nk}^2 = (c_{nk} + c_{kn})^2 + 2\sum_{\nu} c_{nk} c_{\nu n} c_{\nu k} + 2\sum_{\nu} c_{kn} c_{\nu n} c_{\nu k} + \sum_{\mu\nu} c_{\mu n} c_{\mu k} c_{\nu n} c_{\nu k}.$$

Write

$$(12) \qquad s = \sum_{nk} (c_{nk} + c_{kn})^2;$$

the convergence of (12) follows from the convergence of (9) and the general inequality

$$(13) \qquad (a+b)^2 \leq 2a^2 + 2b^2.$$

The series $\sum_{\nu nk} c_{nk} c_{\nu n} c_{\nu k}$ converges absolutely; for, by the convergence of (9), $|c_{nk}|$ remains less than some positive M, we have

$$|c_{nk} c_{\nu n} c_{\nu k}| < M |c_{\nu n} c_{\nu k}|,$$

and we know that (8) converges. Then from (4) we have

$$\sum_{\nu nk} c_{nk} c_{\nu n} c_{\nu k} = \sum_{\nu n} c_{\nu n} \sum_{k} c_{nk} c_{\nu k} = -\sum_{\nu n} c_{\nu n} (c_{\nu n} + c_{n\nu}).$$

Similarly we have

$$\sum_{\nu nk} c_{kn} c_{\nu n} c_{\nu k} = \sum_{\nu k} c_{\nu k} \sum_{n} c_{kn} c_{\nu n} = -\sum_{\nu k} c_{\nu k} (c_{\nu k} + c_{k\nu}) = -\sum_{\nu n} c_{n\nu} (c_{n\nu} + c_{\nu n});$$

the last step is simply a change of notation. By (12) we now find

$$(14) \qquad \sum_{\nu nk} c_{nk} c_{\nu n} c_{\nu k} + \sum_{\nu nk} c_{kn} c_{\nu n} c_{\nu k} = -s.$$

The quadruple series $\sum_{\mu\nu nk} |c_{\mu n} c_{\mu k} c_{\nu n} c_{\nu k}|$ converges; for, by the convergence of (8), $\sum_{\mu} |c_{\mu n} c_{\mu k}|$ for all n and k remains less than some positive M_1, we have

$$\sum_{\mu} |c_{\mu n} c_{\mu k} c_{\nu n} c_{\nu k}| < M_1 |c_{\nu n} c_{\nu k}|,$$

and $\sum_{\nu nk} |c_{\nu n} c_{\nu k}|$ is (8) and known to be convergent. A further use of (4) gives the result

$$(15)\quad \begin{aligned}\sum_{\mu\nu nk} c_{\mu n} c_{\mu k} c_{\nu n} c_{\nu k} &= \sum_{\mu\nu n} c_{\mu n} c_{\nu n} \sum_{k} c_{\mu k} c_{\nu k} = -\sum_{\mu\nu n} c_{\mu n} c_{\nu n} (c_{\mu\nu} + c_{\nu\mu}) \\ &= -\sum_{\mu\nu} (c_{\mu\nu} + c_{\nu\mu}) \sum_{n} c_{\mu n} c_{\nu n} = \sum_{\mu\nu} (c_{\mu\nu} + c_{\nu\mu})^2 = s.\end{aligned}$$

From (11), (12), (14) and (15) we conclude

$$\sum_{nk} \epsilon_{nk}^2 = s - 2s + s = 0,$$

so that every ϵ_{nk} vanishes and Lemma I is established.

5. We shall apply Lemma I to prove

LEMMA II. *If $\{u_n(x)\}$ and $\{\bar{u}_n(x)\}$ are two sets of functions normal and orthogonal on the interval $0 \leq x \leq 1$, if the former set is uniformly bounded on this interval, and if*

$$(16)\quad \bar{u}_n(x) - u_n(x) = \sum_{k} c_{nk} u_k(x), \qquad (n = 1, 2, 3, \ldots),$$

where

$$c_{nk} = \int_0^1 (\bar{u}_n(x) - u_n(x))\, u_k(x)\, dx,$$

and where the series

$$(5)\quad \sum_{n} \Big(\sum_{k} |c_{nk}|\Big)^2$$

converges, then the latter set of functions is also uniformly bounded on the interval and we have the developments

$$(17)\quad u_n(x) - \bar{u}_n(x) = \sum_{k} c_{kn} \bar{u}_k(x), \qquad (n = 1, 2, 3, \ldots).$$

The two series (16) and (17) converge absolutely and uniformly on the interval.

The series $\sum_k |c_{nk}|$ converges and is uniformly bounded for all n by the convergence of (5), so the series (16) converges absolutely and uniformly and the set $\{\bar{u}_n\}$ is uniformly bounded on the interval.

In the normality-orthogonality condition

$$\int_0^1 \bar{u}_n(x)\,\bar{u}_k(x)\,dx = \begin{cases} 1, & n = k, \\ 0, & n \neq k, \end{cases}$$

we insert the absolutely and uniformly convergent series (16) for the $\{\bar{u}_n\}$, multiply out and integrate term by term, observing that the $\{u_n\}$ are normal and orthogonal. The result is (4), and (6) follows by Lemma I. From the convergence of (8) and the properties of $\sum_k |c_{nk}|$ follow the convergence and the uniform boundedness for all n of $\sum_k |c_{nk}|$. Thus the right-hand member of (17) converges absolutely and uniformly.

From the relation (6) multiplied by u_k and summed for all k we have, making use of (16) and the convergence of $\sum_{\nu k} c_{\nu n}\, c_{\nu k}$ proved from the convergence of (8),

$$\sum_k c_{kn}\, u_k + \sum_k c_{nk}\, u_k + \sum_\nu c_{\nu n} \sum_k c_{\nu k}\, u_k = 0,$$

$$\sum_k c_{kn}\, u_k + (\bar{u}_n - u_n) + \sum_\nu c_{\nu n}\,(\bar{u}_\nu - u_\nu) = 0;$$

this last equation becomes essentially (17) if we identify ν with k.

6. We are now in a position to prove our principal theorem.

THEOREM. *Let $\{u_n(x)\}$ and $\{\bar{u}_n(x)\}$ be two sets of functions normal and orthogonal on the interval $0 \leq x < 1$, the former set uniformly bounded on this interval, and such that*

$$(16) \qquad \bar{u}_n(x) - u_n(x) = \sum_{k=1}^{\infty} c_{nk}\, u_k(x), \qquad (n = 1, 2, 3, \ldots),$$

where

$$c_{nk} = \int_0^1 [\bar{u}_n(x) - u_n(x)]\, u_k(x)\, dx$$

and where the series

$$(5) \qquad \sum_{n=1}^{\infty} \left\{ \sum_{k=1}^{\infty} |c_{nk}| \right\}^2$$

converges. Then, if $f(x)$ is any function integrable and with an integrable square, $0 < x < 1$, the two series

$$(18)\qquad f(x) \backsim \sum_{k=1}^{\infty} a_k\, u_k(x),$$

$$(19)\qquad f(x) \backsim \sum_{k=1}^{\infty} b_k\, \bar{u}_k(x)$$

where

$$a_k = \int_0^1 f(x)\, u_k(x)\, dx, \qquad b_k = \int_0^1 f(x)\, \bar{u}_k(x)\, dx,$$

have essentially the same convergence properties; this in the sense that the series corresponding to their term-by-term difference

$$(20)\qquad \sum_{k=1}^{\infty} (a_k\, u_k(x) - b_k\, \bar{u}_k(x))$$

converges absolutely and uniformly on the entire interval to the sum zero.

The sign $\backsim$ is used simply to denote that the coefficients a_k and b_k are given by the formulas indicated. Of course, if $f(x)$ is equal to the series (18), for example, and if we are at liberty to multiply both members of the equation by $u_k(x)$ and to integrate term by term the resulting series, a_k must be given by the formula indicated.

We shall use the two relations obtained by term-by-term integration of (16) and (17) after multiplication by $f(x)\,dx$; this formal work is justified by the convergence of $\sum_k |c_{nk}|$ and $\sum_k |c_{kn}|$,

$$(21)\qquad b_n - a_n = \sum_k c_{nk}\, a_k.$$

$$(22)\qquad a_n - b_n = \sum_k c_{kn}\, b_k.$$

By (22) and (16), (20) is equivalent to

$$(23)\quad \sum_k [(a_k - b_k)\, u_k - b_k\, (\bar{u}_k - u_k)] = \sum_k [\sum c_{rk}\, b_r\, u_k - \sum_r c_{kr}\, b_k\, u_r].$$

The convergence of $\sum_n b_n^2$ (and similarly of $\sum_n a_n^2$) follows from the Bessel inequality

$$(24)\qquad \int_0^1 f^2(x)\, dx \geq \sum_n b_n^2,$$

which may be proved by the relation

$$\int_0^1 \left[f(x) - \sum_{n=1}^{N} b_n \bar{u}_n(x) \right]^2 dx \to 0.$$

The convergence of

$$(25) \qquad \sum_n |b_n| \sum_k |c_{nk}|$$

now follows from (7) and the convergence of (5), and proves the absolute and uniform convergence of the right-hand member of (23) and the theorem.

7. It follows immediately that *properties of absolute convergence, convergence, summability, and divergence at any given point obtain for one of the series* (18), (19) *as for the other, and likewise the properties of uniform convergence in the entire interval considered or in any sub-interval, uniform summability, and also term-by-term integrability. Whenever the two series are convergent, summable, or properly divergent, their sums are the same.* The nature of the approximating functions and of their approach to the limit (in the case of convergence) at a point of continuity or of discontinuity of $f(x)$ is essentially the same for (18) as for (19). In particular, if Gibbs' phenomenon occurs for (18), it also occurs for (19). The reader will notice various other properties* common to the sets $\{u_n\}$ and $\{\bar{u}_n\}$, such as the existence or non-existence of a continuous function for which the formal series does not converge at every point.

If there exists no continuous function $f(x)$, not identically zero, such that all the $\{a_k\}$ are zero, the set $\{u_k\}$ is said to be *closed* with respect to continuous functions.† We have assumed nothing of the closure of $\{u_k\}$ or $\{\bar{u}_k\}$, but it results from (21) and (22) that, if either set is closed, so is also the other.

8. The closure of the set of characteristic functions of the system (1), (2) has recently been proved by Professor Birkhoff‡ from the closure of the set for the system (1), (2) when $g(x) = 0$. It was Professor Birkhoff's note which suggested to me the possibility of the present treatment.

* Thus the simultaneous convergence or divergence of $\sum_n |a_n|$ and $\sum_n |b_n|$ follows from (22) and the convergence of (25).

† There is a corresponding definition and similarity of the property for $\{u_k\}$ and $\{\bar{u}_k\}$ for functions $f(x)$ integrable and with an integrable square.

‡ Proc. Nat. Acad. Sci., vol. 3 (1917), pp. 656–659. The necessary facts concerning the asymptotic nature of the characteristic functions were later proved in detail by Murray, these Annals, ser. (2), vol. 22 (1920–1921), pp. 145–156. The latter paper contains a far more detailed investigation than is necessary for the application of our general theorem.

The methods and general theorem of the present paper are similar to the methods and general theorem of a recent paper* in which there is considered the generalization of a normal and orthogonal system $\{u_k\}$ to a system $\{\bar{u}_k\}$ which is not necessarily normal and orthogonal.

Application to Sturm-Liouville Series.

9. We now apply our main theorem (§ 6) to the case where

$$u_k(x) = \sqrt{2}\sin k\pi x \tag{26}$$

and $\bar{u}_k(x)$ is the k^{th} characteristic function of the system (1), (2). We assume for convenience that $g(x)$ is continuous† on the interval $0 \leqq x \leqq 1$.

We notice directly from the differential equation (2) that $\bar{u}_k(x)$ is a function with a continuous second derivative, so the most elementary theory of Fourier's series informs us that the development (16) is valid. It remains merely to prove the convergence of (5); for this proof we shall need only those most simple asymptotic formulas for $\bar{u}_k(x)$ and ϱ_k (the k^{th} characteristic number of the system (1), (2)) which are readily proved by the original method of Liouville‡:

$$\varrho_k - k\pi \to 0 \text{ as } k \to \infty, \tag{27}$$

$$\bar{u}_k(x) = \sqrt{2}\left(\sin k\pi x + \frac{1}{k}\, q_k(x)\right), \qquad |q_k(x)| < c, \tag{28}$$

for all k and all x on the interval. It is interesting to note that Haar needs the expansion of ϱ_k and $\bar{u}_k(x)$ up to a remainder term of the order $1/n^2$.

The notation for c_{nk} gives us the formula

$$c_{nk} + \delta_{nk} = \int_0^1 \bar{u}_n(x) \cdot \sqrt{2}\sin k\pi x\,dx, \qquad \delta_{nk} = \begin{cases} 1, & n = k, \\ 0, & n \neq k. \end{cases} \tag{29}$$

* Walsh, A generalization of the Fourier cosine series, Trans. Amer. Math. Soc., vol. 22 (1921), pp. 230-239.

† Even this restriction may be lightened with little difficulty. Haar (l. c.) uses the asymptotic expansion given by Hobson, and Hobson supposes $g(x)$ to be of bounded variation.

‡ See, for example, Kneser, Integralgleichungen, pp. 95-98.

If we integrate twice by parts and make use of the boundary conditions (2) for $\bar{u}_n(x)$, we have

$$c_{nk} + \delta_{nk} = -\frac{1}{k^2\pi^2}\int_0^1 \bar{u}_n''(x)\cdot\sqrt{2}\sin k\pi x\,dx. \tag{30}$$

By (27) there exists a number N such that

$$|\varrho_n - n\pi| < \frac{1}{2} \quad \text{for } n > N. \tag{31}$$

We have by (30) [c_1, c_2, etc. denote positive constants],

$$|c_{nk}| < \frac{c_1}{k^2} \quad \text{for } k \neq n, \quad n \leq N,$$

$$|c_{nn}| < 1 + \frac{c_1}{n^2} \quad \text{for } n \leq N;$$

the boundedness of $\bar{u}_n''(x)$ follows from the boundedness of the $\{\bar{u}_n(x)\}$ and equation (1). Hence $\sum\limits_k |c_{nk}|$ converges for every n, and

$$\sum_{n\leq N}\left\{\sum_k |c_{nk}|\right\}^2$$

has a finite value.

10. To prove the convergence of

$$\sum_{n>N}\left\{\sum_k |c_{nk}|\right\}^2,$$

we make $k = n$ in (29) and make use of (28):

$$c_{nn} + 1 = 1 + \int_0^1 \frac{2}{n}\varphi_n(x)\sin n\pi x\,dx,$$

so that

$$|c_{nn}| < \frac{c_2}{n}. \tag{32}$$

For $k \neq n$, we substitute the value of $\bar{u}_n''(x)$, found from the differential equation (1), in (30), whence

$$c_{nk} = \frac{\varrho_n^2}{k^2\pi^2}\int_0^1 \bar{u}_n(x)\sqrt{2}\sin k\pi x\,dx - \frac{1}{k^2\pi^2}\int_0^1 g(x)\bar{u}_n(x)\sqrt{2}\sin k\pi x\,dx.$$

The first integral on the right is precisely c_{nk}, by (29), so we have

$$(33) \qquad (\varrho_n^2 - k^2\pi^2)\,c_{nk} = \int_0^1 g(x)\,\bar{u}_n(x)\sqrt{2}\sin k\pi x\,dx.$$

We are supposing $n > N$ and $k \neq n$, so that $|n-k| \geq 1$, and we have by (31)

$$|\varrho_n - k\pi| = |\varrho_n - n\pi + (n-k)\pi| \geqq |n-k|\pi - |\varrho_n - n\pi| > |n-k|\pi - \frac{1}{2} > |n-k|,$$

$$|\varrho_n + k\pi| = |\varrho_n - n\pi + (n+k)\pi| \geqq (n+k)\pi - |\varrho_n - n\pi| > n+k.$$

Then (33) gives the inequality

$$(34) \qquad |c_{nk}| < \frac{1}{|k^2-n^2|}\left|\int_0^1 g(x)\,\bar{u}_n(x)\sqrt{2}\sin k\pi x\,dx\right|.$$

11. From the identity

$$\frac{1}{k^2-n^2} = \frac{1}{2n}\,\frac{1}{k-n} - \frac{1}{2n}\,\frac{1}{k+n}$$

we find by (13) that

$$\frac{1}{(k^2-n^2)^2} < \frac{1}{2n^2}\,\frac{1}{(k-n)^2} + \frac{1}{2n^2}\,\frac{1}{(k+n)^2}.$$

Moreover we have

$$\sum_k{}' \frac{1}{(k-n)^2} = \sum_{k=1}^{n-1}\frac{1}{(n-k)^2} + \sum_{k=n+1}^{\infty}\frac{1}{(k-n)^2} < 2\sum_k \frac{1}{k^2},$$

$$\sum_k{}' \frac{1}{(k+n)^2} < \sum_k \frac{1}{k^2},$$

where the accent on the summation sign indicates that the term for $k = n$ is to be omitted. We have finally the inequality

$$\sum_k' \frac{1}{(k^2-n^2)^2} < \frac{3}{2n^2} \sum_k \frac{1}{k^2} = \frac{c_3}{n^2}. \tag{35}$$

12. Let us now apply Bessel's inequality (24) to the function $g(x)\,\overline{u}_n(x)$; we obtain

$$\sum_k' \left(\int_0^1 g(x)\,\overline{u}_n(x) \sqrt{2} \sin k\pi x \, dx \right)^2 \leqq \int_0^1 g^2(x)\,\overline{u}_n^2(x)\,dx < c_4. \tag{36}$$

Apply inequality (7) to the series in (35) and (36); we have by (34) that

$$\left(\sum_k' |c_{nk}| \right)^2 < \frac{c_3 c_4}{n^2}. \tag{37}$$

Apply inequality (13) to (32) and the series of (37); we have

$$\left(\sum_k |c_{nk}| \right)^2 < \frac{2c_2^2 + 2c_3 c_4}{n^2} = \frac{c_5}{n^2}, \qquad n > N,$$

whence the convergence of

$$\sum_{n>N} \left(\sum_k |c_{nk}| \right)^2$$

follows immediately.

THE FUNCTIONAL EQUATION $g(x^2) = 2\alpha x + [g(x)]^2$.

By J. H. M. Wedderburn.

1. **Introduction.** The functional equation considered in this paper arose out of an extension of a problem in arrangements which occurs in the theory of linear algebras. In an algebra which is neither associative nor commutative, n factors may be associated in a number of different ways; thus for four factors a_1, a_2, a_3, a_4 we have five different associations in which the subscripts all occur in their natural order, namely

$$a_1\,(a_2 \cdot a_3\,a_4),\qquad (a_1\,a_2)\,(a_3\,a_4),\qquad (a_1 \cdot a_2\,a_3)\,a_4,$$
$$a_1\,(a_2\,a_3 \cdot a_4),\qquad\qquad\qquad (a_1\,a_2 \cdot a_3)\,a_4.$$

The first problem then is to determine the number N_n of such *types* of association for n factors.

It is easily seen* that we can count the number of different types by taking first those in which the left-hand factor consists of one element and the right-hand one of $n-1$, then those in which the first has two elements and the second $n-2$, and so on. Hence

$$N_n = N_1\,N_{n-1} + N_2\,N_{n-2} + \cdots + N_{n-1}\,N_1.$$

If we set

$$f(x) = N_1 x + N_2 x^2 + N_3 x^3 + \cdots,$$

we have, since $N_1 = 1$,

$$f^2(x) = f(x) - x,$$

and therefore, since $f(x)$ vanishes for $x = 0$,

$$f(x) = \frac{1 - \sqrt{1-4x}}{2} = \sum_1^\infty \frac{(2n-2)!}{(n-1)!\,n!}\,x^n,$$

so that

$$N_n = \frac{(2n-2)!}{(n-1)!\,n!}.$$

* This solution (for an equivalent problem) is also given by P. Quarra, Torino Atti vol. 53, (1918) pp. 1044–1047. See also P. Franklin, Question 2681, Amer. Math. Monthly vol. 25 (1918), p. 118 and solution by C. F. Gummer, ibid., vol. 26 (1919) pp. 127–128.

If we assume that multiplication is commutative, the problem of counting types is much more difficult as is seen from the fact that for $n = 4$

$$a_1 (a_2 \cdot a_3 a_4), \qquad (a_1 a_2)(a_3 a_4)$$

are the only types since $a_1 \cdot a_2 a_3 = a_2 a_3 \cdot a_1$ are now of the same type. If n is odd, the method of counting used above is still valid and leads to

$$(1) \qquad 2 N_n = N_1 N_{n-1} + N_2 N_{n-1} + \cdots + N_{n-1} N_1,$$

since the types corresponding to the term $N_{n-r} N_r$ are the same as those belonging to the term $N_r N_{n-r}$. If, however, n is even, say $n = 2k$, the number of types in which the number of elements in the left and right factors is $k = n/2$ is obviously not $N_k^2/2$ but $N_k(N_k + 1)/2$, so that we have in place of (1)

$$(2) \qquad 2 N_n = N_1 N_{n-1} + N_2 N_{n-2} + \cdots + N_{n-1} N_1 + N_{\frac{n}{2}}.$$

If, on the analogy of the solution in the previous problem, we set

$$g(x) = 1 - \sum_1^\infty N_n x^n,$$

we readily derive from (1) and (2) that $g(x)$ satisfies the functional equation

$$g(x^2) = 2x + g^2(x).$$

This equation may be replaced by another of somewhat simpler form by setting

$$(3) \qquad h(x) = g(x)/x^{\frac{1}{2}},$$

which gives

$$h(x^2) = 2 + h^2(x),$$

from which it is obvious that $\varphi_z = h(x^{2^z})$ is a solution of the difference equation

$$\varphi_{z+1} = 2 + \varphi_z^2.$$

This suggests the consideration of the difference equation

$$(4) \qquad \psi_{z+1} = a\psi_z^2 + b\psi_z + c$$

where a, b and c are constants and $a \neq 0$. On setting

$$\varphi_z = a\psi_z + \frac{b}{2}, \qquad \alpha = \frac{4ac+2b-b^2}{8},$$

equation (4) becomes

$$\varphi_{z+1} = 2\alpha + \varphi_z^2,$$

which, as above, leads to the consideration of the functional equation

$$g(x^2) = 2\alpha x + g^2(x) \tag{5}$$

with which we shall be mainly concerned here. The associated function defined by (3) then satisfies the equation

$$h(x^2) = 2\alpha + h^2(x). \tag{6}$$

When it is desired to indicate expressly the dependence of these functions on α, we shall write $g(x, \alpha)$ and $h(x, \alpha)$ in place of $g(x)$ and $h(x)$.

Equations (5) and (6) belong to a class of functional equations considered by Poincaré and Picard. The reader is referred to a series of papers by Fatou* where full references are given.

2. **Solutions which are regular at the origin.** If $g(x)$ is regular at the origin, the value a_0 of $g(0)$ satisfies the equation $a_0^2 = a_0$, so that it is either 0 or 1. If $a_0 = 0$, then also $\alpha = 0$, and it is readily seen that $g(x) = x^m$, m arbitrary. Excluding this exceptional case, we may therefore set

$$g(x) = 1 - a_1 x - a_2 x^2 - \cdots = 1 - \sum_1^\infty a_n x^n \tag{7}$$

in (5) and, comparing coefficients,

$$\begin{aligned}
a_1 &= \alpha \\
2a_2 &= a_1 a_1 + a_1 \\
2a_3 &= a_1 a_2 + a_2 a_1 \\
2a_4 &= a_1 a_3 + a_2 a_2 + a_3 a_1 + a_2 \\
&\cdots\cdots\cdots\cdots\cdots\cdots \\
2a_{2k} &= a_1 a_{2k-1} + a_2 a_{2k-2} + \cdots + a_{2k-1} a_1 + a_k \\
2a_{2k+1} &= a_1 a_{2k} \quad + a_2 a_{2k-1} + \cdots + a_{2k} a_1,
\end{aligned} \tag{8}$$

* P. Fatou. Sur les équations fonctionnelles, Bull. Soc. Math. de France, vol. 47, pp. 161–271, vol. 48, pp. 33–94, 208–384.

which may be written

$$(8') \qquad a_1 = \alpha, \qquad 2\,a_m = \sum_{r=1}^{m-1} a_r\, a_{m-r} + a_{m/2}, \qquad (m = 2, 3, \ldots)$$

if we agree to reckon $a_{m/2}$ as zero when $m/2$ is not an integer.

There are two exceptional values of α in which the solution (7) is trivial; firstly $\alpha = 0$, which gives $g(x) \equiv 1$, and secondly $\alpha = -1$, in which case all the a's vanish except the first and

$$g(x) = 1 + x, \qquad (\alpha = -1).$$

The case $\alpha = 0$ is excluded from further consideration unless specially mentioned and, of course, in the case $\alpha = -1$ the discussion in the remaining sections is trivial.

3. The convergence of the series for $g(x)$. If the term $a_{m/2}$ in (8′) is suppressed, equation (5) becomes

$$(9) \qquad g^2(x) = -2\,\alpha x + 1,$$

or

$$g(x) = \sqrt{1 - 2\,\alpha x} = 1 - \alpha x - \tfrac{1}{2}\,\alpha^2 x^2 - \cdots$$
$$= 1 - \sum_{1}^{\infty} b_n\, x^n;$$

this series converges for $|x| < 1/2|\alpha|$. If we set $\beta_n = |b_n|/2$, then, since each b is the product of a power of α and a positive numerical coefficient, we see readily from (9) that

$$(10) \qquad \begin{aligned} \beta_1 &= |\alpha|/2 \\ \beta_2 &= \beta_1\, \beta_1 \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \beta_m &= \beta_1\, \beta_{m-1} + \beta_2\, \beta_{m-2} + \cdots + \beta_{m-1}\, \beta_1 \\ &\cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \end{aligned}$$

and conversely (10) leads to a convergent series.

If we set $|a_n| = \alpha_n$, then from (8)

$$(11) \qquad \alpha_1 = |\alpha|, \quad 2\,\alpha_m \leqq \alpha_1\, \alpha_{m-1} + \alpha_2\, \alpha_{m-2} + \cdots + \alpha_{m-1}\, \alpha_1 + \alpha_{m/2},$$

and, if γ_n is the sequence of positive numbers defined by

$$\gamma_1 = |\alpha|, \quad 2\gamma_m = \gamma_1\gamma_{m-1} + \gamma_2\gamma_{m-2} + \cdots + \gamma_{m-2}\gamma_1 + \gamma_{m/2}$$

then

$$\alpha_r \leqq \gamma_r, \quad (r = 1, 2, \ldots). \tag{12}$$

For, assuming that this inequality holds for $r < m$, as is certainly the case for $m = 2$, then from (11)

$$2\alpha_m \leqq \sum_1^{m-1} \alpha_r \alpha_{m-r} + \alpha_{m/2} \leqq \sum_1^{m-1} \gamma_r\gamma_{m-r} + \gamma_{m/2} = 2\gamma_m,$$

so that (12) follows by induction. The series (7) for $g(x, \alpha)$ therefore converges if $\sum \gamma_n x^n = g(x, |\alpha|)$ converges.

Let δ be a positive quantity satisfying the conditions

$$\delta > |\alpha|, \qquad \delta > 1,$$

and let δ_m be the sequence of increasing positive numbers defined by

$$\delta_1 = \delta, \qquad \delta_m = \sum_1^{m-1} \delta_r\,\delta_{m-r}, \qquad (m = 2, 3, \ldots),$$

then

$$\gamma_m \leqq \delta_m.$$

For this inequality is true for $m = 1$ and, assuming that it is true for $r < m$, we have

$$2\gamma_m = \sum_1^{m-1} \gamma_r\gamma_{m-r} + \gamma_{m/2} \leqq \sum_1^{m-1} \delta_r\;\delta_{m-r} + \delta_{m/2} = \delta_m + \delta_{m/2} < 2\delta_m,$$

since the sequence of δ's continually increases. Now we have already seen in (10) that $\sum \delta_r x^r$ converges for $|x| < 1/4\delta$; hence $\sum \gamma_n x^n$ converges, and therefore also $\sum \alpha_n x^n$, that is to say, the series (7) converges absolutely for

$$|x| < 1/4|\alpha|, \qquad |x| < 1/4.$$

We have therefore proved that there always exists a unique solution of (5) which is regular at the origin.

When α is real and positive, an upper limit for the radius of convergence may be found as follows. The first three terms in the series for $g(x)$ are

$$1 - \alpha x - \frac{\alpha(\alpha+1)}{2} x^2,$$

the remaining terms all having negative coefficients when $\alpha > 0$; hence $g(\alpha^{-1})$ is negative, if the series converges for that value of x. Now $g(0) = 1 > 0$, so that, if the radius of convergence, R, is greater than $1/\alpha$, there is some value ζ of x for which $g(\zeta) = 0$; and, if we put $x = \zeta^{\frac{1}{2}}$ in (5), we have

$$0 = g(\zeta) = 2\alpha\zeta^{\frac{1}{2}} + g^2(\zeta^{\frac{1}{2}}),$$

so that $g(\zeta^{\frac{1}{2}})$ is imaginary. Since this is impossible so long as both ζ and $\zeta^{\frac{1}{2}}$ are inside the circle of convergence, R is certainly less than the smaller of $1/\alpha$ and $1/\alpha^{\frac{1}{2}}$. In particular, if $\alpha > 1$, then $R < 1$. Closer limits are of course obtained by taking more terms of the series. For instance, if $\alpha = 1$ and $x = 2/3$, the sum of the first three terms of the series is negative and therefore the same argument as before shows that $R < 2/3$. Similarly if $\alpha > (\sqrt{13} - 3)/2 = 0.3027 \ldots$, the radius R is less than 1.

4. **The singularities of** $g(x)$. If we write (5) in the form

$$g(x) = 2\alpha x^{\frac{1}{2}} + g^2\left(x^{\frac{1}{2}}\right), \tag{13}$$

we see immediately that, if $x = \zeta$ is a singularity, so is also $x = \zeta^{\frac{1}{2}}$; and similarly all the points $\zeta^{2^{-n}}$ $(n = 1, 2, \ldots)$ are singularities. If $|\zeta| = 1$, all these points lie on the circle* C_1, while, if $|\zeta| \neq 1$, the points approach more and more closely to this circle as n increases and at the same time become more and more numerous, since there are 2^n determinations of $\zeta^{2^{-n}}$, and moreover in such a way that every point of C_1 is a limit point of the set of singularities. The circle C_1 therefore forms a natural boundary across which $g(x)$ cannot be continued analytically. Hence *the radius of convergence of the series* (7) *is never greater than* 1 *unless it is infinite.*†

In exactly the same way, using (5) in place of (13), it follows that, if ζ is a singularity, so is also ζ^2 unless $g^2(x)$ is regular at $x = \zeta$ which is then a branch point of order 2 at which $g(x) = 0$. If ζ is a singularity for which

---

* The circle with center at the origin and radius $|x| = r$ will be denoted by C_x or C_r.

† It is shown below that the radius is only infinite when $\alpha = -1$ and $g(x) = 1 + x$.

$|\zeta|$ is a minimum and we assume $|\zeta| < 1$, then $|\zeta| > 0$ and, since $|\zeta^2| < |\zeta|$, it follows that $g^2(x)$ is regular at $x = \zeta$, which is therefore a branch point of order 2 at which $g(x) = 0$. We may therefore set

$$g(x) = (x - \zeta)^{\frac{1}{2}} g_1(x - \zeta),$$

where $g_1(x)$ is regular at $x = \zeta$. Moreover, if $|\zeta|$ is not a minimum but is still less than 1, then, since $|\zeta^n| \to 0$ as $n \to \infty$ while $g(x)$ is regular at $x = 0$, $g(x)$ must be regular at ζ^n for some value of n. It then follows from (5) that $g(\zeta)$ is finite and has a finite number of determinations; such singularities are therefore algebraic. We shall now show that, if there is a singularity inside C_1, there is a *unique* singularity for which $|\zeta|$ is a minimum.

We have already seen in (6) that $h(x) = g(x)/x^{\frac{1}{2}}$ satisfies the equation

$$(14) \qquad h(x^2) = 2\alpha + h^2(x).$$

If x_1 and x_2 are two different zeros of $g(x)$ which lie inside C_1, then

$$(15) \qquad h(x_1^{2^n}) = h(x_2^{2^n}),$$

as each is the same polynomial in 2α, e. g., if in (14) $h(x) = 0$, then

$$h(x^2) = 2\alpha,$$

$$h(x^4) = 2\alpha + h^2(x^2) = 2\alpha + 4\alpha^2,$$

and so on. Now $f(x) = 1/h^2(x) = x/g^2(x)$ is regular at $x = 0$, and

$$f'(x) = \frac{g(x) - 2xg'(x)}{g^3(x)},$$

which has the value 1 at $x = 0$. But in (15), $|x_1^{2^n}|$ and $|x_2^{2^n}|$ can be made as small as we please by making n sufficiently large; hence there are an infinity of distinct pairs of points x' and x'' in any neighbourhood of the origin for which

$$\frac{f(x') - f(x'')}{x' - x''} = 0.$$

This is impossible since $f'(0) = 1$; hence $g(x)$ vanishes for at most one value of x within C_1. We have already shown that $g(x)$, and therefore also $h(x)$,

vanishes at the branch point nearest the origin and hence it follows that there is not more than one such point.

We have therefore shown that $g(x)$ *has no singularities within the circle* C_1 *except possibly branch points of finite order; and, if it has any singularity within* C_1, *there is a unique singularity* ζ *for which* $|\zeta|$ *is a minimum; this point is a branch point of order 2 at which* $g(x) = 0$ *and it is the only zero within* C_1. *Every point* $\zeta^{1/2^n}$ *is also a singular point and every singular point within* C_1 *is of this form.*

Exactly the same argument as above may be used to show that $h(x)$ never takes on the same value twice within C_1.

We have seen above that C_1 in general forms a natural boundary for $g(x)$. There exist, however, solutions which have a simple pole at $x = \infty$ but otherwise behave in the region exterior to C_1 in much the same way as $g(x)$ does in the interior region, or, more explicitly, if $\zeta_1, \zeta_2, \ldots$ are the singularities of $g(x)$ within C_1, then

$$\bar{g}(x) = xg\left(\frac{1}{x}\right)$$

is a solution of (5) whose only singularities outside C_1 are $\zeta_1^{-1}, \zeta_2^{-1}, \ldots$ and a simple pole at $x = \infty$. For, if $|x| > |\zeta^{-1}|$, ζ being as before the singularity of $g(x)$ whose modulus is least, then

$$\bar{g}(x^2) = x^2 g\left(\frac{1}{x^2}\right) = x^2\left[\frac{2\alpha}{x} + g^2\left(\frac{1}{x}\right)\right]$$

$$= 2\alpha x + \left[xg\left(\frac{1}{x}\right)\right]^2 = 2\alpha x + \bar{g}^2(x).$$

In particular, when $g(x) = 1 + x$, we have $\bar{g}(x) \equiv g(x)$.

It is readily seen that no entire function except $1 + x$ can be a solution of (5). For differentiating the terms of this equation we have

$$g(x^2) = 2\alpha x + [g(x)]^2$$

$$2xg'(x^2) = 2\alpha + 2gg',$$

$$4x^2 g''(x^2) + 2g'(x^2) = 2gg'' + 2g'^2$$

$$\cdots\cdots\cdots\cdots\cdots\cdots$$

$$2^n x^n g^{(n)}(x^2) + \frac{n(n-1)}{2} 2^{n-1} x^{n-2} g^{(n-1)}(x^2) + \cdots$$

$$= 2g(x)\, g^{(n)}(x) + 2ng'(x)\, g^{(n-1)}(x) + \cdots$$

Hence the value of $g^{(n)}(1)$ is determined uniquely in terms of $g(1)$ unless $g(1) = 2^{n-1}$, i. e., $2\alpha = 2^{n-1} - 2^{2n-2}$ for any positive integral value of n. For these values of α, we have, on putting $x = -1$ in (5),

$$g(1) = -2^{n-1} + 2^{2n-2} + [g(-1)]^2$$

or

$$[g(-1)]^2 = 2^{n-1} + 2^{n-1} - 2^{2n-2} = -(2^{2n-2} - 2^n) < 0$$

unless $n = 1, 2$. Since the series for $g(x)$ has real values when α and x are real, $g(-1)$ cannot have a negative square so that we need only consider the cases $n = 1, 2$. For $n = 2$, we have $\alpha = -1$, and, as we have already seen, $g(x) = 1 + x$ which is an entire function; for $n = 1$, $\alpha = 0$ and $g(x) = 1$, again an entire function; we may therefore exclude these two trivial cases.

Now $\bar{g}(x)$ satisfies the same functional equation as $g(x)$ and $\bar{g}(1) = g(1)$, and therefore the derivatives of $\bar{g}(x)$ have the same values at $x = 1$ as those of $g(x)$. But, if $g(x)$ is an entire function, $\bar{g}(x)$ is regular at $x = 1$; it follows therefore that the series defining $g(x)$ and $\bar{g}(x)$ are identical. This is impossible as $g(x)$ is regular at $x = 0$ while $\bar{g}(x)$ has a singularity there except for the cases $\alpha = 0,1$ considered above. We have therefore shown that $g(x)$ *is not an entire function except when* $\alpha = 0$, $g(x) = 1$, *and* $\alpha = -1$, $g(x) = 1 + x$.

5. **The radius of convergence.** The radius of convergence of (7) can be calculated by means of (14) under certain restrictions and, when α is real, also the value of x which corresponds to a given real value of $g(x)$.

By successive applications of (14) we have

$$(16)\qquad \begin{aligned} h(x^2) &= 2\alpha + h^2(x) \\ h(x^4) &= 2\alpha + (2\alpha + h^2(x))^2 \\ &\cdots\cdots\cdots\cdots \\ h(x^{2^n}) &= 2\alpha + (2\alpha + \cdots + (2\alpha + h^2(x))^2 \cdots)^2, \end{aligned}$$

where in $h(x^{2^n}) = g(x^{2^n})/x^{2^{n-1}}$ the term 2α occurs n times. If we set

$$(17)\qquad k_n(x) = |h(x^{2^n})|^{\frac{1}{2^{n-1}}} = \frac{|g(x^{2^n})|^{\frac{1}{2^{n-1}}}}{|x|},$$

then, since, if $|x| < 1$, $x^{2^n} \to 0$ as $n \to \infty$ and $g(0) = 1$, it follows that

$$\lim_{n \to \infty} k_n(x) = \frac{1}{|x|}. \tag{17'}$$

If we replace $h(x)$ by λ in (16) and write*

$$\begin{array}{llll} & h_0 = \lambda, & & k_0 = |\lambda^2|, \\ & h_1 = 2\alpha + \lambda^2, & & k_1 = |2\alpha + \lambda^2|, \\ (18) & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot & & \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ & h_n = 2\alpha + h_{n-1}^2, & & k_n = |h_n|^{\frac{1}{2^{n-1}}}, \end{array}$$

then k_n approaches a definite limit as $n \to \infty$ provided λ is admissable as a value of $h(x)$ for $|x| < 1$. The convergence of k_n to its limit is, as a rule, rapid for real values of α which are not too small, and this furnishes a practical method of calculating the zero $x = \zeta$, which corresponds to $\lambda = 0$, provided we are able otherwise to determine that it lies inside C_1.

The values of ζ calculated in this manner for certain values af α are

α	ζ	α	ζ
0.25	0.9292	5.00	0.09533
0.50	0.6654	10.00	0.04880
1.00	0.4027	50.00	0.00995
2.00	0.2230		

The convergence of k_n to its limit can also be determined independently under certain conditions. We shall suppose in the first case that for some value of r

$$|h_r| \overline{\overline{>}} 1 + |2\alpha|^{\frac{1}{2}} \tag{19}$$

so that

$$|h_r|^2 \overline{\overline{>}} 1 + 2|2\alpha|^{\frac{1}{2}} + |2\alpha| > |2\alpha|;$$

then

$$|h_{r+1}| = |2\alpha + h_r^2| > |h_r|^2 - |2\alpha| \overline{\overline{>}} 1 + 2|2\alpha|^{\frac{1}{2}},$$

* When it is necessary to indicate the values of λ and 2α explicitly, we shall write $h_n(\lambda, 2\alpha)$ in place of h_n.

whence it follows easily by induction that

(20) $$|h_{r+n}| \overline{\overline{>}} 1 + 2^n |2\alpha|^{\frac{1}{2}};$$

and therefore if n is sufficiently large,* $|h_n| > |2\alpha|$.
Now

$$\frac{k_n}{k_{n-1}} = \left|\frac{h_n}{h_{n-1}^2}\right|^{\frac{1}{2^{n-1}}} = \left|1 + \frac{2\alpha}{h_{n-1}^2}\right|^{\frac{1}{2^{n-1}}} \leqq \left(1 + \frac{|2\alpha|}{|h_{n-1}^2|}\right)^{\frac{1}{2^{n-1}}}$$

$$\leqq (1 + |2\alpha|)^{\frac{1}{2^{n-1}}} < 1 + \frac{|2\alpha|}{2^{n-1}}$$

for n sufficiently large. The infinite product $\prod k_n/k_{n-1}$ therefore converges and so k_n approaches a definite, finite limit.

If $|2\alpha| > 2.25$, it is easily shown that (19) is satisfied for $\lambda = 0$ and $r = 2$.

The convergence follows in the same way if it is known that $|h_n| \overline{\overline{>}} \epsilon > 0$ for all values of $n > r$, ϵ being independent of n.

If $\lambda = 0$ is a possible value of λ and z, $|z| < 1$, the corresponding value of x, then as before

$$h(z^2) = 2\alpha,$$

$$h(z^4) = 2\alpha + (2\alpha)^2$$

$$\cdots\cdots\cdots\cdots$$

$$h(z^{2^n}) = 2\alpha + h^2(z^{2^{n-1}}).$$

Hence $h(z^{2^n})$ is a polynomial in α of degree 2^{n-1} whose coefficients are positive integers which do not depend on z or α; this polynomial we shall denote by $p_n(2\alpha)$. If $\mu = 2\alpha$ is a root of $p_n(\mu) = 0$, then $h(z^{2^n}) = 0$, so that z^{2^n} is a root of $g(x)$. Hence $(z^{2^n})^{2^n}$ is also a root, and so on. This, however, is impossible if $|z| < 1$, as the origin would then be an essential singularity of $g(x)$; we can therefore conclude that, *if 2α is a root of any of the polynomials $p_n(\mu)$, $g(x)$ vanishes at no point within C_1 which is therefore the circle of convergence of the series* (7), except in the two trivial cases in which this series represents an entire function.

6. **The polynomials $h_n(\lambda, \mu)$ and $p_n(\mu)$.** The polynomial $h_n(\lambda, \mu)$ is the polynomial defined in (18) with μ in place of 2α, i.e.,

(21) $$h_0(\lambda, \mu) = \lambda, \qquad h_n(\lambda, \mu) = \mu + h_{n-1}^2(\lambda, \mu),$$

---

* If $|h_n| > 1 + |2\alpha|^{\frac{1}{2}}$ for every n greater than a certain value, then evidently $\lim k_n \leqq 1$.

while $p_n(\mu) = h_n(0, \mu)$ so that

$$(21') \qquad p_1(\mu) = \mu, \qquad p_n(\mu) = \mu + p_{n-1}^2(\mu).$$

The following properties follow readily from these definitions.

In the first place we have

$$(22) \qquad h_n(\lambda, \mu) = h_{n-r}(h_r(\lambda, \mu), \mu) = p_{n-r}(\mu) + q_{nr}(\lambda, \mu)\, h_r^2(\lambda, \mu),$$

where $q_{nr}(\lambda, \mu)$ is a polynomial in λ, μ with positive integral coefficients; hence in particular

$$(22') \qquad p_n(\mu) = p_{n-r}(\mu) + q_{nr}(0, \mu)\, p_r^2(\mu).$$

From these equations we deduce immediately that

(i) *$h_n(\lambda, \mu)$ and $h_r(\lambda, \mu)$ have no common factor; and if s is the H. C. F. of n and r, $p_s(\mu)$ is the H. C. F. of $p_n(\mu)$ and $p_r(\mu)$.*

To prove the first part of this lemma we need only observe that from (22) every common factor of h_n and h_r is a factor of $p_{n-r}(\mu)$ and therefore does not contain λ. This is, however, impossible since the coefficient of the highest power of λ in h_r is unity.

To prove the second part we observe from (22′) that every common factor of p_n and p_r is a factor of p_{n-r}; and hence, by a repetition of this argument, it is a factor of p_s. It only remains, therefore, to prove that p_s is a factor of p_{ks}, k being any integer. Putting $n = 2s, 3s, \ldots, r = s, 2s, \ldots$ in (22′), we have

$$p_{2s} = p_s + q_{2s,s}\, p_s^2, \qquad p_{3s} = p_{2s} + q_{3s,s}\, p_s^2,$$

and so on; from which the required result follows immediately. This also shows that $p_1 = \mu$ is a factor of every $p_n(\mu)$ as is of course obvious otherwise.

(ii)

$$(23) \quad h_n(\lambda, \mu) - h_r(\lambda, \mu) = (h_{n-1} + h_{r-1})(h_{n-2} + h_{r-2}) \cdots (h_{n-r} + h_0)(h_{n-r} - h_0),$$

$$(23') \quad p_n(\mu) - p_r(\mu) = (p_{n-1} + p_{r-1})(p_{n-2} + p_{r-2}) \cdots (p_{n-r+1} + p_1)\, p_{n-r}^2.$$

For

$$h_s(\lambda, \mu) - h_t(\lambda, \mu) = h_{s-1}^2(\lambda, \mu) - h_{t-1}^2(\lambda, \mu)$$

$$= [h_{s-1}(\lambda, \mu) + h_{t-1}(\lambda, \mu)]\, [h_{s-1}(\lambda, \mu) - h_{t-1}(\lambda, \mu)].$$

An immediate consequence of (23) is that $h_n(\lambda, \mu)$, $(n = k, k+1, k+2, \ldots)$ all have the same value for values of λ and μ for which

$$h_r(\lambda, \mu) + h_{r-1}(\lambda, \mu) = 0, \qquad (r = 1, 2, \cdots k),$$

an important particular case of which is

$$p_n(-2) = 2, \qquad (n \geqq 2).$$

(iii) Differentiating (21) with regard to λ^2 and μ, we readily prove by induction that

(24) $$\frac{\partial h_n}{\partial \mu} = 1 + 2\, h_{n-1} + 2^2 h_{n-1}\, h_{n-2} + 2^3 h_{n-1}\, h_{n-2}\, h_{n-3} + \cdots$$
$$\cdots + 2^{n-1}\, h_{n-1}\, h_{n-2} \cdots h_1,$$

(24′) $$\frac{\partial h_n}{\partial \lambda^2} = 2^{n-1}\, h_{n-1}\, h_{n-2} \cdots h_1,$$

an interesting particular case of which is $p'_n(-2) = -(2^{2n-1}+1)/3$, $(n \geqq 2)$.

(iv) *If*

(25) $$|\lambda^2 + \mu| \geqq \tfrac{1}{2} + \sqrt{|\mu| + \tfrac{1}{4}},$$

then

(26) $$|h_n(\lambda, \mu)| \geqq |\lambda^2 + \mu|^2 - |\mu| \geqq |\mu|^{\frac{1}{2}}, \qquad (n \geqq 2),$$

and, in particular, if $|\mu| \geqq 2$,

(26′) $$|p_n(\mu)| \geqq |\mu|^2 - |\mu| \geqq |\mu|^{\frac{1}{2}}, \qquad (n \geqq 2).$$

If (25) is satisfied, then

$$|\mu| + |\mu|^{1/2} < |\mu| + \tfrac{1}{2} + \sqrt{|\mu| + \tfrac{1}{4}} = (\tfrac{1}{2} + \sqrt{|\mu| + \tfrac{1}{4}})^2 \geqq 1;$$

hence

$$|\lambda^2 + \mu|^2 \geqq |\mu| + |\mu|^{1/2},$$

or

$$|\lambda^2 + \mu|^2 - |\mu| \geqq |\mu|^{1/2} \geqq 0.$$

Now, for $n = 2$,

$$|h_2(\lambda, \mu)| = |\mu + (\lambda^2 + \mu)^2| \geqq |\lambda^2 + \mu|^2 - |\mu| > |\mu|^{1/2}.$$

Let us assume, therefore, that (26) is true for $2, 3, \ldots n$; then

$$|h_{n+1}(\lambda, \mu)| = |h_n^2(\lambda, \mu) + \mu| > |h_n^2(\lambda, \mu)| - |\mu|$$

$$> |\lambda^2 + \mu|^4 - 2|\mu|\,|\lambda^2 + \mu|^2 + |\mu|^2 - |\mu|$$

$$= |\lambda^2 + \mu|^2 - |\mu| + [\,|\lambda^2 + \mu|^2 - (|\mu| + \tfrac{1}{2})]^2 - |\mu| - \tfrac{1}{4}$$

$$> |\lambda^2 + \mu|^2 - |\mu| > |\mu|^{1/2}$$

by (25). Equation (26) then follows by induction.

If $\lambda = 0$, (25) becomes

$$|\mu| \geqq \tfrac{1}{2} + \sqrt{|\mu| + \tfrac{1}{4}},$$

or $|\mu|^2 > 2|\mu|$, i. e., $|\mu| > 2$; and (26) becomes (26′).

Since $|\lambda^2 + \mu| \geqq |\,|\lambda|^2 - |\mu|\,|$, (25) is satisfied by

$$|\lambda|^2 > |\mu| + \tfrac{1}{2} + \sqrt{|\mu| + \tfrac{1}{4}}, \qquad |\lambda|^2 > |\mu|,$$

or

$$|\lambda|^2 \leq |\mu| - \tfrac{1}{2} - \sqrt{|\mu| + \tfrac{1}{4}}, \qquad |\lambda|^2 < |\mu| \geq 2.$$

(v) It follows immediately from (26′) that *the absolute value of every root of $p_n(\mu)$, except $\mu = 0$, is less than 2 and therefore corresponds to a value of $|\alpha| < 1$ and, if α is real, to a value of α between -1 and 0.*

(vi) *If μ_n is the real negative root of p_n of greatest absolute value, there is one, and only one, real root μ_{n+1} of p_{n+1} between -2 and μ_n.*

When $\mu = -2$, we have already seen that $p_n = 2 > 0$, and if $\mu = \mu_{n-1}$, then $p_n = \mu_{n-1} + p_{n-1}^2 = \mu_{n-1} < 0$; there is therefore at least one real root of p_n between -2 and μ_{n-1}.

If we differentiate (21′) twice, we get

$$p_n''(\mu) = 2[p_{n-1}'(\mu)]^2 + 2p_{n-1}(\mu)\,p_{n-1}''(\mu),$$

so that, if μ is real, p_n'' is positive if p_{n-1} and p_{n-1}'' have the same sign. For $n=3$, $\mu_2=-1$; and, for $\mu<-1$, we evidently have $p_3''>0$, so that to the left of $\mu=\mu_3$ both p_3 and p_3'' are positive. It follows that, to the left of μ_3, $p_4''>0$; and so on. We can therefore conclude that p_n has one, and only one, real root between -2 and μ_{n-1} as otherwise p_n'', which is equal to or greater than 0, would change sign for some value of μ between these limits.

It is also easy to show that $\mu_n \to -2$ as $n\to\infty$. For

$$p_n(-2)=2, \qquad p_n(\mu_{n-1})=\mu_{n-1}<-1,$$

while $p_n''(\mu)>0$ between -2 and μ_{n-1} so that the graph of $p_n(\mu)$ is concave upwards between these limits; μ_n therefore lies to the left of the line joining $(-2,2)$ to $(\mu_{n-1},-1)$, whence

$$\mu_{n-1}-\mu_n>\tfrac{1}{3}(\mu_{n-1}+2).$$

(vii) *If μ and λ are real and positive, and if $\mu<\frac{1}{4}$ and $\mu+\lambda^2-\lambda<0$ or $\lambda<\mu+\lambda^2<\frac{1}{2}$, then*

$$\lim_{n=\infty} h_n(\lambda,\mu)=\tfrac{1}{2}-\sqrt{\tfrac{1}{4}-\mu}.$$

For all other positive values of μ and λ, $h_n(\lambda,\mu)\to\infty$ as $n\to\infty$.

From (23) with $r=n-1$ we have

$$h_n-h_{n-1}=(h_{n-1}+h_{n-2})(h_{n-2}+h_{n-3})\cdots(h_1+h_0)(\mu+\lambda^2-\lambda).$$

If $\mu+\lambda^2-\lambda<0$, which requires $\mu<1/4$, the h's therefore form a decreasing sequence of positive quantities and so approach a finite limit. If $\mu+\lambda^2-\lambda>0$, they form an increasing sequence. If the sequence of h's has a finite limit, then, since $h_n=\mu+h_{n-1}^2$, we must have $l=\mu+l^2$. Moreover, since $h_{n-1}+h_{n-2}\to 2l$, we see immediately from (23) that $l<1/2$, and hence $l=\frac{1}{2}-\sqrt{\frac{1}{4}-\mu}$. A finite limit can therefore only exist if $\mu<1/4$.

If $\mu<1/4$ and $h_1=\mu+\lambda^2+1/2$, then

$$h_2=\mu+(\mu+\lambda^2)^2<\tfrac{1}{2},$$

and therefore, by an easy induction, $h_n<1/2$ for every n. On the other hand, if $\lambda<\mu+\lambda^2>1/2$, the h's form an increasing sequence which cannot have a finite limit since $h_{n-1}+h_{n-2}$ is greater than unity.

(viii) *If* $\mu > 1/4$, $\lambda > 0$ *and* $\frac{1}{x} = \lim_{n=\infty} k_n(\lambda, \mu)$, *then* $\lambda = h(x, \mu)$.

In the first place, $k(\lambda, \mu) = \lim_{n=\infty} k_n(\lambda, \mu)$ exists and is finite. For, under the given conditions, h_n increases indefinitely with n and hence the condition of (19) and (20) of § 5 that $h_n > 1 + 2^{n-r}\mu^{\frac{1}{2}}$ is satisfied for n greater than some finite value of r. The value of the limit is moreover greater than unity; for we can write

$$h_n > (1 + \epsilon)^{2n-1}, \qquad (\epsilon > 0)$$

for some n sufficiently large and ϵ sufficiently small, whence

$$h_{n+1} > \mu + (1 + \epsilon)^{2n} > (1 + \epsilon)^{2n},$$

so that the inequality also holds, with the same ϵ, for every subsequent value of n.

We shall now show that $\partial k_n / \partial \lambda$ approaches a finite value as $n \to \infty$. From (17) and (24′) we have

$$\frac{1}{2\lambda}\frac{\partial k_n}{\partial \lambda} = \frac{\partial k_n}{\partial \lambda^2} = k_n \frac{h_{n-1} h_{n-2} \cdots h_1}{h_n} = k_n(\lambda)\, q_n(\lambda)$$

say. Now

$$\frac{q_{n+1}}{q_n} = \frac{h_n^2}{h_{n+1}} = \frac{1}{1 + \frac{\mu}{h_n^2}},$$

and $\sum \mu / h_n^2$ converges since $h_n > 1 + 2^{n-r}\mu^{\frac{1}{2}}$ for n sufficiently large. Hence q_n approaches a definite finite limit as $n \to \infty$, and this limit is never zero. Moreover, k_n and q_n approach their limits uniformly as regards λ since in both cases the convergence was obtained by comparison with series which are independent of λ. Hence

$$\frac{\partial k(\lambda, \mu)}{\partial \lambda} = 2\lambda k(\lambda, \mu)\, q(\lambda, \mu).$$

It follows that, in any interval $0 < \epsilon \leq \lambda \leq N$ in which ϵ is as small and N as great as we please, $x = 1/k(\lambda, \mu)$ is a uniformly continuous function of λ and

possesses a derivative which is nowhere zero. There exists therefore a unique single-valued inverse function $\lambda = H(x, \mu)$. Now we saw in § 5 that $k(\lambda, \mu) = 1/x$ if λ is admissable as a value of $h(x, \mu)$; also $h(x, \mu) \to \infty$ as $x \to 0$; hence, if N is taken large enough, the range of values for $h(x, \mu)$ will overlap that for $H(x, \mu)$, and in the common part of their ranges these two functions have the same value. Both functions satisfy the same functional equation, namely $h(x^2) = \mu + h^2(x)$; hence, if a is so small that a and $a^{1/2}$ both lie in the common range, so will also $a^{1/4}$ provided always that $h(x, \mu)$ and $H(x, \mu)$ do not vanish; and so on. The two functions therefore have the same range of definition and this range is from ϵ to the point at which they vanish or, if they do not vanish, from ϵ to 1. We can then use (17′) to calculate the zero of $g(x)$ provided $\mu > 1/4$, and *the radius of convergence of $g(x)$ is less than unity for real values of $2\alpha \equiv \mu > 1/4$, and it is equal to unity for* $0 < \mu \leq 1/4$.

7. **Solutions which have a singular point at the origin.** If $g(x)$ has a pole of order n at $x = 0$, then

$$G(x) = x^n g(x)$$

is regular there and $G(0) \neq 0$. The function $G(x)$ satisfies the functional equation

$$G(x^2) = 2\alpha x^{2n+1} + G^2(x), \tag{27}$$

so that $G(0) = 1$. Since $2n + 1$ is integral whenever n is an integral multiple of $1/2$, the discussion of this equation will also embrace those solutions which have an algebroid pole or zero whose order is of the form $m/2$.

In place of (27) we shall consider the more general equation

$$F(x^2) = 2(\alpha^0 + \alpha' x + \alpha'' x^2 + \cdots + \alpha^{(n)} x^n) + F^2(x), \tag{28}$$

and we shall only consider solutions which are regular at the $x = 0$. Substituting in this equation

$$F(x) = a_0 - a_1 x - a_2 x^2 - a_3 x^3 - \cdots \tag{29}$$

we obtain

$$a_0^2 = a_0 - 2\alpha^0$$

$$2a_0 a_1 = 2\alpha'$$

$$2a_0 a_2 = a_1 a_1 + a_1 + 2\alpha''$$

$$\cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot$$

$$2a_0 a_n = \sum_{r=1}^{n-1} a_r a_{n-r} + a_{n/2} + 2\alpha^n$$

$$2a_0 a_m = \sum_{r=1}^{m-1} a_r a_{m-r} + a_{m/2} \qquad (m > n).$$

If $a_0 = 0$, then $\alpha^0 = 0 = \alpha'$, so that $F_1(x) = F(x)/x$ is regular at $x = 0$ and satisfies the equation

$$F_1(x^2) = 2(\alpha'' + \alpha''' x + \cdots + \alpha^{(n)} x^{n-2}) + F_1^2(x);$$

we may therefore assume $a_0 \neq 0$ without real loss of generality.

The discussion of the convergence of (29) is very similar to that already given in § 3 and is therefore given here merely in outline. Let

$$a_r = a_0 b_r, \qquad |b_r| = \beta_r, \qquad |a_0| = \alpha_0, \qquad |\alpha^{(r)}| = \alpha_0^2 \beta^{(r)},$$

then

$$2\beta_m \leqq \sum \beta_r \beta_{m-r} + \frac{1}{\alpha_0}\beta_{m/2} + 2\beta^{(m)},$$

the last term being absent if $m > n$. There are two cases to consider.

(i) Suppose $1/\alpha_0 \leqq 1$. Let

$$\gamma_1 = \gamma, \qquad 2\gamma_m = \sum \gamma_r \gamma_{m-r} + \gamma_{m/2},$$

where γ is chosen so large that $\beta_m \leq \gamma_m$ $(m = 1, 2, \ldots n)$. If this inequality holds for all values of the subscript from 1 to $m - 1 \geqq n$, then

$$2\beta_m \leq \sum \gamma_m \gamma_{m-r} + \frac{1}{\alpha_0}\gamma_{m/2} \leq \sum \gamma_r \gamma_{m-r} + \gamma_{m/2} = 2\gamma_m,$$

so that it also holds for every value of m. The proof that the series converges is then exactly as in § 3.

(ii) Suppose that $1/\alpha_0 > 1$. Then, γ being chosen as before, let

$$\gamma_1 = \gamma, \qquad 2\gamma_m = \sum \gamma_r \gamma_{m-r} + \frac{1}{\alpha_0}\gamma_{m/2},$$

so that

$$2\beta_m \leq \sum \gamma_r \gamma_{m-r} + \frac{1}{\alpha_0}\gamma_{m/2} = 2\gamma_m, \qquad (m > n).$$

Let δ be a positive quantity satisfying the conditions

$$\delta \geqq \alpha_0 \gamma, \qquad \delta \geqq 1,$$

and set $\delta_1 = \delta$, $\delta_m = \sum \delta_r \delta_{m-r}$; then $\delta_m \geqq \alpha_0^m \gamma_m$. This inequality is true when $m = 1$; suppose it is true for all values of the subscript up to $m-1$, then

$$2\gamma_m = \sum \gamma_r \gamma_{m-r} + \frac{1}{\alpha_0}\gamma_{m/2} \leqq \sum \left(\frac{1}{\alpha_0}\right)^m \delta_r \delta_{m-r} + \left(\frac{1}{\alpha_0}\right)^{\frac{m+2}{2}} \delta_{m/2}$$

$$= \left(\frac{1}{\alpha_0}\right)^m \delta_m + \left(\frac{1}{\alpha_0}\right)^{\frac{m+2}{2}} \delta_{m/2}.$$

But $\delta_{m/2} \leqq \delta_m$ and $m > (m+2)/2$; hence the inequality holds for every value of m. It then follows exactly as before that (29) converges for $|x| < \alpha_0/4\delta$.

It is not difficult to show* that no solution of (29) can be an entire function unless it is a polynomial. It then follows that the original functional equation (5) cannot have a solution of the form

$$g(x) = a_m x^m + \cdots + a_0 + \sum_1^\infty a_{-r} x^{-r}$$

in which the infinite series does not terminate. For, if there were such a solution, $G\left(\frac{1}{x}\right) = g(x)/x^m$ would be an entire function of $z = 1/x$ which

* One method of proof is to consider the manner in which $|F(x)|$ increases with $|x|$.

satisfies the equation

$$G(z^2) = 2\alpha z^{2m-1} + G^2(z)$$

and is not a polynomial.

8. **Conclusion.** The method of calculating the zeros of $g(x)$ which was given in § 5 may also be used to calculate the value of $g(x)$ for any real x and α. When this is carried out, it is immediately seen that comparatively few steps of the limiting process are required to give a fairly accurate result. If the number of steps required for a given degree of accuracy has been ascertained, the process may by profitably inverted, e. g., if four steps are sufficient, we may set

$$h(x) = [[[[x^{-8} - 2\alpha]^{\frac{1}{2}} - 2\alpha]^{\frac{1}{2}} - 2\alpha]^{\frac{1}{2}} - 2\alpha]^{\frac{1}{2}},$$

or in general

$$h(x) = [[\cdots[[x^{-2^{n-1}} - 2\alpha]^{\frac{1}{2}} - 2\alpha]^{\frac{1}{2}} \cdots - 2\alpha]^{\frac{1}{2}} - 2\alpha]^{\frac{1}{2}},$$

where the square root is extracted n times. If n is taken so large that $[x^{-2^{n-1}} - 2\alpha]^{\frac{1}{2}} = x^{-2^{n-2}}$ to the degree of accuracy required, then nothing is gained by increasing the number of steps beyond n.

PRINCETON, N. J.,
December, 1921.

ON CONVERGENCE FACTORS IN TRIPLE SERIES AND THE TRIPLE FOURIER'S SERIES.

BY BESS M. EVERSULL.

Although the triple Fourier's series has been used somewhat extensively, no proof of the validity of the development of a function of three variables in such a series has been given. It is the purpose of this paper to establish certain facts in connection with the summability of such a series, with a view to applying the theory of convergence factors to some problems in mathematical physics. As an illustration of such applications it will be shown that the formal series arising from the discussion of a certain problem in the flow of heat furnishes an actual solution to the problem. A method of dealing with problems of this type has been originated by Fejér* and applied by him to problems involving the ordinary Fourier's series; the same method has been applied to problems involving the double Fourier's series by Professor C. N. Moore† and will here be applied to the problem involving the triple Fourier's series which we wish to consider.

1. **Summability of triple series.** The type of summability with which this paper is concerned will first be defined. Consider the triple series

$$(1) \qquad \sum_{l=1,\, m=1,\, n=1}^{\infty,\, \infty,\, \infty} a_{lmn}$$

and form

(2)

$$S_{lmn}^{(r)} = \sum_{i=1,\, j=1,\, k=1}^{l,\, m,\, n} \frac{\Gamma(r+l-i)}{\Gamma(r)\Gamma(l-i+1)} \cdot \frac{\Gamma(r+m-j)}{\Gamma(r)\Gamma(m-j+1)} \cdot \frac{\Gamma(r+n-k)}{\Gamma(r)\Gamma(n-k+1)}\, s_{ijk},$$

where

$$(3) \qquad s_{ijk} = \sum_{p=1,\, q=1,\, s=1}^{i,\, j,\, k} a_{pqs}.$$

If the quotient

$$(4) \qquad \frac{S_{lmn}^{(r)}}{A_{lmn}^{(r)}},$$

---

* Cf. his paper *Untersuchungen über Fouriersche Reihen*, Math. Annalen, vol. 58 (1903–1904), p. 51.

† See his paper *On convergence factors in double series and the double Fourier's series*, Trans. Amer. Math. Soc., vol. 14 (1913), pp. 73–104. The summability of the double Fourier's series has also been considered by W. H. Young (Proc. Lond. Math. Soc., ser. 2, vol. 11 (1912), p. 133), and by W. W. Küstermann (cf. his dissertation, *Über Fouriersche Doppelreihen und das Poissonsche Doppelintegral*; Munich, 1913).

where

$$(5)\qquad A^{(r)}_{lmn} = \frac{\Gamma(l+r)}{\Gamma(r+1)\Gamma(l)} \cdot \frac{\Gamma(m+r)}{\Gamma(r+1)\Gamma(m)} \cdot \frac{\Gamma(n+r)}{\Gamma(r+1)\Gamma(n)},$$

approaches a limit S as l, m and n become infinite, the triple series is said to be summable (Cr) and to have a value equal to this limit. This type of summability for a triple series is analogous to that which Cesàro considered for an ordinary series. The above definition is valid for any r, real or complex, except zero or a negative integer. We may include the case where $r=0$, if we assume that the right hand sides of equations (2) and (5) have the values they approach as r approaches zero. In this case, summability ($C0$) is the same as convergence as defined by Pringsheim. For the applications we wish to make, we need only consider the case where r is zero or a positive integer.

Before going on with our investigation, it will be advantageous to develop some properties of $S^{(r)}_{lmn}$ and $A^{(r)}_{lmn}$. From the definitions (2) and (5), it is evident that

$$(6)\qquad \sum_{l=1,\, m=1,\, n=1}^{\infty\ \infty\ \infty} S^{(r)}_{lmn} x^l y^m z^n \backsim (1-x)^{-r}(1-y)^{-r}(1-z)^{-r} \sum_{l=1,\, m=1,\, n=1}^{\infty\ \infty\ \infty} s_{lmn} x^l y^m z^n.$$

$$(7)\qquad \sum_{l=1,\, m=1,\, n=1}^{\infty\ \infty\ \infty} A^{(r)}_{lmn} x^l y^m z^n \backsim xyz(1-x)^{-(r+1)}(1-y)^{-(r+1)}(1-z)^{-(r+1)},$$

the sign $\backsim$ indicating that, if the expressions on each side of it are expanded in ascending powers of x, y and z, the coefficients of the corresponding terms will be equal.

We may also derive the relation*

$$(8)\qquad a_{lmn} = \sum_{i=0,\, j=0,\, k=0}^{r+1,\, r+1,\, r+1} (-1)^i(-1)^j(-1)^k \binom{r+1}{r+1-i}\binom{r+1}{r+1-j} \cdot \binom{r+1}{r+1-k} S^{(r)}_{l-i,m-j,n-k},$$

where for the sake of uniformity we set

$$(9)\qquad S^{(r)}_{pqs} = 0 \qquad (p,\ q \text{ or } s \leq 0).$$

---

* The derivation of this expression may be generalized from that for the simple series, given in Bromwich, *Theory of infinite series*, p. 317.

From the definition of $A_{lmn}^{(r)}$, we have

$$\lim_{l,m,n\to\infty} \frac{A_{lmn}^{(r)}}{l^r\, m^r\, n^r} = \left\{\frac{1}{\Gamma(r+1)}\right\}^3,$$

and therefore

(10) $$A_{lmn}^{(r)} < K l^r\, m^r\, n^r \qquad (l, m, n = 1, 2, 3, \ldots),$$

where K is a positive constant.

In order to make our definition of summability of wide use, it must be consistent with the definition of convergence; that is, if the series is summable $(C0)$, or convergent according to Pringsheim's definition, it should also be summable (Cr) for any $r > 0$, and to the same value. We shall prove that our definition satisfies this requirement for all classes of convergent triple series for which *the condition of finitude** is satisfied, i. e. for which

(11) $$s_{lmn} = \left|\frac{S_{lmn}^{(0)}}{A_{lmn}^{(0)}}\right| < C \qquad (l, m, n = 1, 2, 3, \ldots),$$

C being a positive constant. It is possible that this restriction is more narrow than necessary, but this question will not be considered here as the theorems on convergence factors which we shall have to prove present the same restriction.

If a convergent triple series satisfies the condition (11), it also possesses the property that as l, m and n become infinite, the quotient $S_{lmn}^{(r)}/A_{lmn}^{(r)}$ converges to the same value as s_{lmn}, and will also remain less in absolute value than the positive constant C for all values of l, m and n. Hence, to insure consistency with the definition of convergence, we shall consider only triple series which satisfy the conditions that

(12) $$\lim_{l,m,n\to\infty} \frac{S_{lmn}^{(r)}}{A_{lmn}^{(r)}} \text{ exists,}$$

and

(13) $$\left|\frac{S_{lmn}^{(r)}}{A_{lmn}^{(r)}}\right| < C \qquad (l, m, n = 1, 2, 3, \ldots)$$

where C is a positive integer, and r a positive integer or zero.

* This term was introduced by Bromwich and Hardy in their paper, *Some extensions to multiple series of Abel's Theorem*, Proc. Lond. Math. Soc., ser. 2, vol. 2 (1904), p. 161.

Before proceeding to the discussion of the theorem of consistency, it would be well to note that the $(r+1)^{\text{th}}$ difference of the triple sequence $f_{ijk}\ (i, j, k = 1, 2, 3, \ldots)$ is entirely analogous to the $(r+1)^{\text{th}}$ difference of the simple sequence, and is defined by

$$\text{(14)} \quad \triangle_{r+1}^{r+1}{}_{r+1} f_{ijk} = \sum_{p=0,\, q=0,\, s=0}^{r+1,\, r+1,\, r+1} (-1)^p (-1)^q (-1)^s \cdot \binom{r+1}{r+1-p}\binom{r+1}{r+1-q}\binom{r+1}{r+1-s} f_{i+p,\, j+q,\, k+s}.$$

We shall now proceed to the discussion of the consistency theorem, proving first two necessary lemmas.

LEMMA 1. *If the two triple sequences*

$$\text{(15)} \qquad a_{lmn},\ b_{lmn} \qquad (l, m, n = 1, 2, 3, \ldots)$$

satisfy the conditions

$$\text{(a)} \qquad \left|\frac{\triangle_1^1{}_1\, a_{lmn}}{\triangle_1^1{}_1\, b_{lmn}}\right|^* < C, \qquad (l, m, n = 0, 1, 2, 3, \ldots),$$

$$\text{(b)} \qquad \triangle_1^1{}_1\, b_{lmn} < 0, \qquad (l, m, n = 1, 2, 3, \ldots),$$

$$\text{(c)} \qquad \lim_{l, m, n \to \infty} b_{lmn} = \infty,$$

$$\text{(d)} \qquad \begin{aligned} &\lim_{p \to \infty} \frac{b_{\lambda q s}}{b_{pqs}} = 0, \quad \lim_{q \to \infty} \frac{b_{p\mu s}}{b_{pqs}} = 0, \quad \lim_{s \to \infty} \frac{b_{pq\nu}}{b_{pqs}} = 0, \\ &\lim_{p, q \to \infty} \frac{b_{\lambda\mu s}}{b_{pqs}} = 0, \quad \lim_{p, s \to \infty} \frac{b_{\lambda q\nu}}{b_{pqs}} = 0, \quad \lim_{q, s \to \infty} \frac{b_{p\mu\nu}}{b_{pqs}} = 0, \end{aligned}$$

$$(\lambda, \mu, \nu = 1, 2, 3, \ldots),$$

---

* In forming $\triangle_1^1{}_1\, a_{lmn}$ and $\triangle_1^1{}_1\, b_{lmn}$, for the sake of uniformity we set

$$a_{lmn} = 0 = b_{lmn} \qquad (l, m \text{ or } n \leq 0).$$

where the limits are approached uniformly, the first for all positive integral values of q and s, the second for all positive integral values of p and s, the third of p and q, the fourth of s, the fifth of q and the sixth of p, and

$$\text{(e)} \qquad \lim_{l,m,n\to\infty} \frac{\triangle_{1\,1}^{1} a_{lmn}}{\triangle_{1\,1}^{1} b_{lmn}} \ \textit{exists,}$$

then we shall have

$$\text{(16)} \qquad \left|\frac{a_{lmn}}{b_{lmn}}\right| < C, \qquad (l, m, n = 1, 2, 3, \ldots),$$

and

$$\text{(17)} \qquad \lim_{l,m,n\to\infty} \frac{a_{lmn}}{b_{lmn}}$$

will exist and be equal to the limit in (e).

We shall first establish the inequality (16). From conditions (a) and (b), we have

$$\text{(18)} \qquad \left|\triangle_{1\,1}^{1} a_{lmn}\right| < -C\triangle_{1\,1}^{1} b_{lmn},$$

and therefore

$$\text{(19)} \qquad \sum_{i=0,j=0,k=0}^{l-1,m-1,n-1} \left|\triangle_{1\,1}^{1} a_{ijk}\right| < -C \sum_{i=0,j=0,k=0}^{l-1,m-1,n-1} \triangle_{1\,1}^{1} b_{ijk}.$$

From this and condition (b),

$$\text{(20)} \qquad \frac{\displaystyle\sum_{i=0,j=0,k=0}^{l-1,m-1,n-1} \left|\triangle_{1\,1}^{1} a_{ijk}\right|}{\displaystyle\sum_{i=0,j=0,k=0}^{l-1,m-1,n-1} \left|\triangle_{1\,1}^{1} b_{ijk}\right|} < C.$$

In view of the fact that $a_{lmn} = 0$ when l, m or $n \leq 0$, it is obvious that

$$\sum_{i=0,j=0,k=0}^{l-1,m-1,n-1} \triangle_{1\,1}^{1} a_{ijk} = -a_{lmn}; \qquad \sum_{i=0,j=0,k=0}^{l-1,m-1,n-1} \triangle_{1\,1}^{1} b_{ijk} = -b_{lmn},$$

and therefore, making use of (20), we have

$$(21)\qquad \left|\frac{a_{lmn}}{b_{lmn}}\right| = \left|\frac{\sum\limits_{i=0,j=0,k=0}^{l-1,m-1,n-1} \Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} a_{ijk}}{\sum\limits_{i=0,j=0,k=0}^{l-1,m-1,n-1} \Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} b_{ijk}}\right| \leqq \frac{\sum\limits_{i=0,j=0,k=0}^{l-1,m-1,n-1} \left|\Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} a_{ijk}\right|}{\sum\limits_{i=0,j=0,k=0}^{l-1,m-1,n-1} \left|\Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} b_{ijk}\right|}.$$

We have thus proved that the relation (16) is true, and it remains only to be shown that the limit (17) exists and is equal to the limit in (e). Since the limit in (e) exists, we can find a λ, μ and ν corresponding to a positive ϵ as small as we please, such that

$$(22)\qquad L-\epsilon < \frac{\Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} a_{lmn}}{\Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} b_{lmn}} < L+\epsilon, \qquad \begin{pmatrix} l \geqq \lambda \\ m \geqq \mu \\ n \geqq \nu \end{pmatrix},$$

where L is the value of the limit in (e). Hence from condition (b) it follows that

$$(23)\qquad (L-\epsilon)\,\Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} b_{lmn} > \Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} a_{lmn} > (L+\epsilon)\,\Delta\begin{smallmatrix}1\\1\\1\end{smallmatrix} b_{lmn} \qquad \begin{pmatrix} l > \lambda \\ m \geqq \mu \\ n > \nu \end{pmatrix}.$$

If we add all the inequalities (23) for all sets of values of l, m and n such that $\lambda \leqq l \leqq p$, $\mu \leqq m \leqq q$, $\nu \leqq n \leqq s$, we have

$$\begin{aligned} &(L-\epsilon)\,(b_{\lambda\mu\nu} - b_{p\mu\nu} - b_{\lambda q\nu} - b_{\lambda\mu s} + b_{pq\nu} + b_{p\mu s} + b_{\lambda qs} - b_{pqs}) \\ &\qquad > a_{\lambda\mu\nu} - a_{p\mu\nu} - a_{\lambda q\nu} - a_{\lambda\mu s} + a_{pq\nu} + a_{p\mu s} + a_{\lambda qs} - a_{pqs} \\ &> (L+\epsilon)\,(b_{\lambda\mu\nu} - b_{p\mu\nu} - b_{\lambda q\nu} - b_{\lambda\mu s} + b_{pq\nu} + b_{p\mu s} + b_{\lambda qs} - b_{pqs}). \end{aligned}$$

From condition (c) it is obvious that we can find a p, q and s so large that $b_{pqs} > 0$, and hence we may divide the above inequality by b_{pqs}, which gives

$$\begin{aligned} &(L-\epsilon)\left[1 - \frac{b_{\lambda qs}}{b_{pqs}} - \frac{b_{p\mu s}}{b_{pqs}} - \frac{b_{pq\nu}}{b_{pqs}} + \frac{b_{\lambda\mu s}}{b_{pqs}} + \frac{b_{\lambda q\nu}}{b_{pqs}} + \frac{b_{p\mu\nu}}{b_{pqs}} + \frac{b_{\lambda\mu\nu}}{b_{pqs}}\right] \\ (24)\qquad &< \frac{a_{pqs}}{b_{pqs}} - \frac{a_{\lambda qs}}{b_{pqs}} - \frac{a_{p\mu s}}{b_{pqs}} - \frac{a_{pq\nu}}{b_{pqs}} + \frac{a_{\lambda\mu s}}{b_{pqs}} + \frac{a_{\lambda q\nu}}{b_{pqs}} + \frac{a_{p\mu\nu}}{b_{pqs}} + \frac{a_{\lambda\mu\nu}}{b_{pqs}} \\ &< (L-\epsilon)\left[1 - \frac{b_{\lambda qs}}{b_{pqs}} - \frac{b_{p\mu s}}{b_{pqs}} - \frac{b_{pq\nu}}{b_{pqs}} + \frac{b_{\lambda\mu s}}{b_{pqs}} + \frac{b_{\lambda q\nu}}{b_{pqs}} + \frac{b_{p\mu\nu}}{b_{pqs}} + \frac{b_{\lambda\mu\nu}}{b_{pqs}}\right]. \end{aligned}$$

From conditions (c) and (d) it follows that, as p, q and s become infinite, the limits of the parentheses in (24) which involve only b's exist and are equal to unity. The last seven terms of the second member of the inequality (24) may be written in the form

$$-\frac{a_{\lambda qs}}{b_{\lambda qs}}\cdot\frac{b_{\lambda qs}}{b_{pqs}}-\frac{a_{p\mu s}}{b_{p\mu s}}\cdot\frac{b_{p\mu s}}{b_{pqs}}-\frac{a_{pq\nu}}{b_{pq\nu}}\cdot\frac{b_{pq\nu}}{b_{pqs}}+\frac{a_{\lambda\mu s}}{b_{\lambda\mu s}}\cdot\frac{b_{\lambda\mu s}}{b_{pqs}}$$

$$+\frac{a_{\lambda q\nu}}{b_{\lambda q\nu}}\cdot\frac{b_{\lambda q\nu}}{b_{pqs}}+\frac{a_{p\mu\nu}}{b_{p\mu\nu}}\cdot\frac{b_{p\mu\nu}}{b_{pqs}}-\frac{a_{\lambda\mu\nu}}{b_{\lambda\mu\nu}}\cdot\frac{b_{\lambda\mu\nu}}{b_{pqs}};$$

hence, by virtue of conditions (c) and (d) and the inequality (21), the limit as p, q and s become infinite, of these seven terms, exists and is equal to zero.

Hence, from (24), we have

$$(25)\qquad L-\epsilon<\underline{\lim}\,\frac{a_{pqs}}{b_{pqs}}\leqq\overline{\lim}\,\frac{a_{pqs}}{b_{pqs}}<L+\epsilon,$$

and since ϵ is an arbitrarily small positive quantity, it follows from (25) that

$$\underline{\lim}\,\frac{a_{pqs}}{b_{pqs}}=\overline{\lim}\,\frac{a_{pqs}}{b_{pqs}}=L,$$

and consequently the limit (17) exists and is equal to L, the limit in (e), and therefore the lemma is proved.

LEMMA 2. *If the triple series* (1) *is summable* (Cr) *where r is zero or a positive integer, and if moreover*

$$(26)\qquad \left|\frac{S^{(r)}_{lmn}}{A^{(r)}_{lmn}}\right|<C,\qquad (l,m,n=1,2,3,\ldots),$$

where C is a positive constant, the series will be summable $(C\,\overline{r+1})$ *to the same value to which it is summable* (Cr), *and furthermore we shall have*

$$(27)\qquad \left|\frac{S^{(r+1)}_{lmn}}{A^{(r+1)}_{lmn}}\right|<C,\qquad (l,m,n=1,2,3,\ldots).$$

From relations (6) and (7) we have

$$\sum_{l=1,\,m=1,\,n=1}^{\infty,\,\infty,\,\infty} S_{lmn}^{(r+1)} x^l y^m z^n \backsim (1+x+x^2+\cdots)(1+y+y^2+\cdots) \cdot (1+z+z^2+\cdots) \sum_{l=1,\,m=1,\,n=1}^{\infty,\,\infty,\,\infty} S_{lmn}^{(r)} x^l y^m z^n,$$

$$\sum_{l=1,\,m=1,\,n=1}^{\infty,\,\infty,\,\infty} A_{lmn}^{(r+1)} x^l y^m z^n \backsim (1+x+x^2+\cdots)(1+y+y^2+\cdots) \cdot (1+z+z^2+\cdots) \sum_{l=1,\,m=1,\,n=1}^{\infty,\,\infty,\,\infty} A_{lmn}^{(r)} x^l y^m z^n.$$

Equating the coefficients of the corresponding terms in each of these expressions, we have

$$(28) \qquad S_{lmn}^{(r+1)} = \sum_{i=1,\,j=1,\,k=1}^{l,\,m,\,n} S_{ijk}^{(r)}; \qquad A_{lmn}^{(r+1)} = \sum_{i=1,\,j=1,\,k=1}^{l,\,m,\,n} A_{ijk}^{(r)},$$

and from these equations we obtain

$$(29) \qquad \triangle\begin{smallmatrix}1\\1\\1\end{smallmatrix} S_{lmn}^{(r+1)} = -S_{l+1,\,m+1,\,n+1}^{(r)}; \qquad \triangle\begin{smallmatrix}1\\1\\1\end{smallmatrix} A_{lmn}^{(r+1)} = -A_{l+1,\,m+1,\,n+1}^{(r)}.$$

We may now apply Lemma 1, taking $S_{lmn}^{(r+1)}$ and $A_{lmn}^{(r+1)}$ as the two triple sequences. It is necessary first to show that the conditions under which Lemma 1 holds true are satisfied. Condition (a) is satisfied; for, from (29),

$$\left|\frac{\triangle\begin{smallmatrix}1\\1\\1\end{smallmatrix} S_{lmn}^{(r+1)}}{\triangle\begin{smallmatrix}1\\1\\1\end{smallmatrix} A_{lmn}^{(r+1)}}\right| = \left|\frac{S_{l+1,\,m+1,\,n+1}^{(r)}}{A_{l+1,\,m+1,\,n+1}^{(r)}}\right|,$$

which from (26) is less than a positive constant C $(l,\ m,\ n\ =\ 1, 2, 3, \ldots)$. Condition (b) is satisfied by virtue of the second of equations (29) and the definition (5). That conditions (c) and (d) are fulfilled also follows from definition (5). That condition (e) is satisfied follows from (29) and the hypothesis of this lemma that the triple series is summable (Cr).

Hence it follows from Lemma 1 that $S_{lmn}^{(r+1)}/A_{lmn}^{(r+1)}$ satisfies the condition of finitude, and that the series is summable $(C\overline{r+1})$ to the same value to which it is summable (Cr).

We may now prove the consistency theorem itself.

THEOREM I. *If the series* (1) *is summable* (Cr), *where* r *is zero or a positive integer, and*

$$\left|\frac{S_{lmn}^{(r)}}{A_{lmn}^{(r)}}\right| < C, \qquad (l, m, n = 1, 2, 3, \ldots),$$

where C *is a positive constant, then the series will be summable* (Cr'), *where* r' *is any integer greater than* r, *to the same value to which it is summable* (Cr), *and furthermore we shall have*

$$\left|\frac{S_{lmn}^{(r')}}{A_{lmn}^{(r')}}\right| < C \qquad (l, m, n = 1, 2, 3, \ldots).$$

If $r' = r + 1$, the theorem reduces to Lemma 2. If $r' > r + 1$, the theorem may be proved by successive applications of Lemma 2.

2. **Convergence factors in triple series.** Before proceeding with the work of this section, it will be necessary to introduce and define a notation which we shall need. If we set

$$(30) \quad \triangle^{\substack{r+1,\,w\\ r+1,\,v\\ r+1,\,u}} = \sum_{p=0,\,q=0,\,s=0}^{r+1-u,\,r+1-v,\,r+1-w} (-1)^p (-1)^q (-1)^s \cdot \binom{r+1}{r+1-p}\binom{r+1}{r+1-q}\binom{r+1}{r+1-s} f_{i+p,\,j+q,\,k+s},$$

we see that it is analogous to the notation (14), and also that the right hand side of (30) is equal to that of (14) with certain of its terms suppressed; the terms which are suppressed are indicated by the indices added to those of expression (14). We may further abbreviate our notation by setting

$$\triangle^{\substack{r+1,\,r+1\\ r+1,\,v\\ r+1,\,u}} = \triangle^{\substack{0\\ r+1,\,v\\ r+1,\,u}}, \text{ etc.}$$

We are now ready to prove the theorems on convergence factors, restricting ourselves in this paper to the treatment of convergence factors in triple series which are summable $(C1)$, as this is the only case necessary for the applications we wish to make.

LEMMA 3. *If the triple series* (1) *is summable* $(C1)$ *and, moreover, the condition* (13) *is satisfied for* $r = 1$, *then, for all positive values of* α, β *and* γ, *the series*

$$(31) \qquad \sum_{l=1,\,m=1,\,n=1}^{\infty\,\infty\,\infty} a_{lmn} f_{lmn}(\alpha, \beta, \gamma)$$

will converge and have the same value as the series

$$\sum_{i=1,j=1,k=1}^{\infty\,\infty\,\infty} S_{ijk}^{(1)} \Delta_{2}^{\frac{2}{2}} f_{ijk}(\alpha, \beta, \gamma), \tag{32}$$

which will also be convergent, provided the convergence factors $f_{ijk}(\alpha, \beta, \gamma)$ *satisfy the conditions*

$$\text{(a)} \qquad \sum_{i=1,j=1,k=1}^{\infty,\infty,\infty} ijk \left| \Delta_{2}^{\frac{2}{2}} f_{ijk}(\alpha, \beta, \gamma) \right| < K, \qquad (\alpha, \beta, \gamma > 0),$$

$$\text{(b)} \qquad \lim_{l\to\infty} l \sum_{j=1,k=1}^{\infty,\infty} jk \left| f_{ljk}(\alpha, \beta, \gamma) \right| = 0, \qquad (\alpha, \beta, \gamma > 0)$$

and the two other conditions of the same type,

$$\text{(c)} \qquad \lim_{l\to\infty,\, m\to\infty} lm \sum_{k=1}^{\infty} K \left| f_{lmk}(\alpha, \beta, \gamma) \right| = 0, \qquad (\alpha, \beta, \gamma) > 0),$$

and the two other conditions of this type, and

$$\text{(d)} \qquad \lim_{l\to\infty,\, m\to\infty,\, n\to\infty} [l\, m\, n\, f_{lmn}(\alpha, \beta, \gamma)] = 0, \qquad (\alpha, \beta, \gamma > 0),$$

it being understood that the limiting processes indicated by conditions (a) — (d) *all have a meaning.*

Substituting in the expression

$$\sum_{i=1,j=1,k=1}^{l,m,n} a_{ijk} f_{ijk}(\alpha, \beta, \gamma) \tag{33}$$

the value of a_{ijk} given by putting $r = 1$ in (8), and rearranging the terms, we have

$$\sum_{i=1,j=1,k=1}^{l,m,n} a_{ijk} f_{ijk} = \sum_{i=1,j=1,k=1}^{l-2,m-2,n-2} S_{ijk}^{(1)} \Delta_{2}^{\frac{2}{2}} f_{ijk} + R_{ijk}^{(1)}, \tag{34}$$

where $R_{ijk}^{(1)}$ consists of three terms of each of the types

$$\sum_{j=1,k=1}^{m-2,n-2} S_{l-1,j,k}^{(1)} \Delta_{\begin{smallmatrix}2\\2\\2,1\end{smallmatrix}} f_{l-1,j,k}, \quad \sum_{j=1,k=1}^{m-2,n-2} S_{ljk}^{(1)} \Delta_{\begin{smallmatrix}2\\2\\0\end{smallmatrix}} f_{ljk}, \quad \sum_{k=1}^{n-2} S_{l-1,m-1,k}^{(1)} \Delta_{\begin{smallmatrix}2\\2,1\\2,1\end{smallmatrix}} f_{l-1,m-1,k},$$

$$\sum_{k=1}^{n-2} S_{l-1,m,k}^{(1)} \Delta_{\begin{smallmatrix}2\\0\\2,1\end{smallmatrix}} f_{l-1,m,k}, \quad \sum_{k=1}^{n-2} S_{l,m-1,k}^{(1)} \Delta_{\begin{smallmatrix}2\\2,1\\0\end{smallmatrix}} f_{l,m-1,k},$$

$$\sum_{k=1}^{n-2} S_{lmk}^{(1)} \Delta_{\begin{smallmatrix}2\\0\\0\end{smallmatrix}} f_{lmk}, \quad S_{l-1,m-1,n}^{(1)} \Delta_{\begin{smallmatrix}0\\2,1\\2,1\end{smallmatrix}} f_{l-1,m-1,n}, \quad S_{l-1,m,n}^{(1)} \Delta_{\begin{smallmatrix}0\\0\\2,1\end{smallmatrix}} f_{l-1,m,n-1},$$

and in addition the terms $S_{l-1,m-1,n-1}^{(1)} \Delta_{\begin{smallmatrix}2,1\\2,1\\2,1\end{smallmatrix}} f_{l-1,m-1,n-1}$ and $S_{lmn}^{(1)} f_{lmn}$.

We shall now show that the right-hand side of (34) approaches a limit as l, m and n become infinite for all positive values of α, β and γ, and that the second term approaches the limit zero. Then it will follow that the left-hand side approaches a limit under the same conditions, and that these limits will be equal. Hence we may conclude that the series (31) converges to the same value as the series (32), the value of (32) being the limiting value of the first term in the expansion (34).

The first term of the right-hand side of (34) is the sum of the $(l-2)(m-2)(n-2)$ terms of the series (32), contained in a rectangular parallelopiped of dimensions $l-2$, $m-2$, $n-2$, taken from the upper, forward, left-hand corner of the series. From condition (13) where $r=1$, we have for the general term of (32)

$$(35) \qquad \left| S_{lmn}^{(1)} \Delta_{\begin{smallmatrix}2\\2\\2\end{smallmatrix}} f_{lmn} \right| < Clmn \left| \Delta_{\begin{smallmatrix}2\\2\\2\end{smallmatrix}} f_{lmn} \right| \qquad \left(\begin{matrix} l, m, n = 1, 2, 3 \ldots \\ \alpha, \beta, \gamma > 0 \end{matrix} \right).$$

From condition (a) it follows that the series whose general term is the right-hand side of (35) converges, and therefore the series (32) is absolutely convergent, and the first term of (34) approaches as a limit the value to which the series (32) converges as l, m and n become infinite.

By expanding each of the terms, except the last, of $R_{ijk}^{(1)}$, and applying conditions (13) for $r=1$, and the appropriate one of the conditions (b)–(d), we may show that these terms approach zero as l, m and n become infinite. That the last term of $R_{ijk}^{(1)}$ approaches zero as l, m and n become infinite may be shown by applying condition (13) for $r=1$, and the condition (d). Hence $R_{ijk}^{(1)}$ approaches the limit zero as l, m and n become infinite, and the lemma is proved.

We are now ready to prove the theorem:

THEOREM II. *If the triple series* (1) *satisfies the conditions of Lemma 3, and the convergence factors* $f_{ijk}(\alpha, \beta, \gamma)$ *satisfy the conditions* (a)—(d) *inclusive of that lemma, and the additional conditions that*

(e) $f_{ijk}(\alpha, \beta, \gamma)$ *is continuous in* α, β *and* γ $\quad\left(\begin{matrix} i, j, k = 1, 2, 3, \ldots \\ \alpha, \beta, \gamma > 0 \end{matrix}\right)$,

(f) $\lim\limits_{\alpha, \beta, \gamma \to 0} [f_{ijk}(\alpha, \beta, \gamma)] = f_{ijk}(0, 0, 0) = 1 \qquad (i, j, k = 1, 2, 3, \ldots)$,

(g) $\lim\limits_{\alpha, \beta, \gamma \to 0} \sum\limits_{i=1}^{i=\infty} i \left| \Delta^{\substack{2\\2}}_{2} f_{ijk}(\alpha, \beta, \gamma) \right| = 0$ *for every j and k, and the two other conditions of the same type,*

(h) $\lim\limits_{\alpha, \beta, \gamma \to 0} \sum\limits_{i=1, j=1}^{i=\infty, j=\infty} i\,j \left| \Delta^{\substack{2\\2}}_{2} f_{ijk}(\alpha, \beta, \gamma) \right| = 0$ *for every k, and the two other conditions of this type, and*

(i) $\sum\limits_{i=1, j=1, k=1}^{\infty, \infty, \infty} i\,j\,k \left| \Delta^{\substack{2\\2}}_{2} f_{ijk} \right|$ *is uniformly convergent in* $(\alpha \geq \alpha_0 > 0,\ \beta \geq \beta_0 > 0,\ \gamma \geq \gamma_0 > 0)$,

then the series (31) *will define a function of* α, β *and* γ, $F(\alpha, \beta, \gamma)$ *which is continuous for all positive values of* α, β *and* γ, *and for which*

$$\lim_{\alpha, \beta, \gamma \to +0} [F(\alpha, \beta, \gamma)] = S, \tag{36}$$

where S *is the value of the series* (1).

From Lemma 3, we know that the series (31) converges to the same value as the series (32) for all positive values of α, β and γ. Hence if we can show that the series (32) defines a function of α, β and γ, which is continuous for all positive values of α, β and γ, and for which equation (36) holds true, our theorem will have been proved.

Since by hypothesis, the series is summable $(C1)$ to S, we may write

$$\frac{S^{(1)}_{lmn}}{l\,m\,n} = S + \epsilon_{lmn} \qquad \left(\lim_{l, m, n \to \infty} \epsilon_{lmn} = 0\right). \tag{37}$$

Using (37), we can reduce the series (32) to the form

$$S \sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} i\,j\,k\, \Delta^{\substack{2\\2}}_{2} f_{ijk}(\alpha, \beta, \gamma) + \sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} i\,j\,k\,\epsilon_{ijk}\, \Delta^{\substack{2\\2}}_{2} f_{ijk}(\alpha, \beta, \gamma). \tag{38}$$

We can evaluate the series in the first term of (38) by applying Lemma 3 to the triple series for which

$$a_{111} = 1, \quad a_{ijk} = 0 \qquad (i, j \text{ or } k > 1), \tag{39}$$

observing that this series is convergent and satisfies the restriction (11). Hence by Theorem 1 this series is summable $(C1)$, and therefore we can say that for this series the equation

$$\sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} a_{ijk} f_{ijk}(\alpha, \beta, \gamma) = \sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} S_{ijk}^{(1)} \triangle_{2}^{\frac{2}{2}} f_{ijk}(\alpha, \beta, \gamma), \tag{40}$$

which expresses the equality between the expressions (31) and (32), reduces to the form

$$f_{111}(\alpha, \beta, \gamma) = \sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} ijk \triangle_{2}^{\frac{2}{2}} f_{ijk}(\alpha, \beta, \gamma). \tag{41}$$

Substituting this value in the expression (38), that expression becomes

$$S f_{111}(\alpha, \beta, \gamma) + \sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} ijk\, \varepsilon_{ijk} \triangle_{2}^{\frac{2}{2}} f_{ijk}(\alpha, \beta, \gamma). \tag{42}$$

The first term is a continuous function of α, β and γ for all positive values of α, β and γ, by virtue of condition (e), and by condition (f) approaches the value S as α, β and γ approach zero from the positive direction. Hence if we can show that the second term of the expression (42), which is equal to the expression (32), is a continuous function of α, β and γ for all positive values of those variables, and approaches zero as α, β and γ approach zero, we shall have proved that the expression (42) is continuous for all positive values of α, β and γ, and approaches S as a limit as α, β and γ approach zero through positive values, and the theorem will have been proved.

By virtue of condition (i) and equations (11) and (37), we infer that the series in the second term of (42) is uniformly convergent in the region

$$(\alpha \geq \alpha_0 > 0, \ \beta \geq \beta_0 > 0, \ \gamma \geq \gamma_0 > 0). \tag{43}$$

From condition (e) its terms are continuous there, and therefore the series will be continuous in the region (43), and consequently, since α_0, β_0 and γ_0 are arbitrary positive quantities, will be continuous for all positive values

of α, β and γ. That it approaches zero as α, β and γ approach zero follows from the fact that ϵ_{ijk} remains finite for all values of i, j and k, and approaches zero as i, j and k become infinite, together with conditions (f), (g) and (h). Therefore, as pointed out above, the theorem has been proved.

The above theorem has been proved only for the case where the series (1) consists of constant terms. By extending the proofs of Lemma 3 and Theorem II, they may be made to apply to the case where the series (1) consists of terms which are functions of three variables. The analogous statement for this case is made in the following corollary:

COROLLARY. *If the triple series* (1), *whose terms are functions of x, y and z, is uniformly summable* ($C1$) *to $f(x, y, z)$ throughout the region R, and*

$$\left| \frac{S^{(1)}_{lmn}(x, y, z)}{l\,m\,n} \right| < C, \qquad \left(\begin{matrix} l, m, n = 1, 2, 3, \dots \\ R \end{matrix} \right),$$

where C is a positive constant, and the convergence factors $f_{ijk}(\alpha, \beta, \gamma)$ satisfy the conditions of Theorem II, the series

$$\sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} a_{ijk}(x, y, z) f_{ijk}(\alpha, \beta, \gamma)$$

will converge uniformly for all positive values of α, β and γ, and for all values of x, y and z in R, and its value, $F(x, y, z, \alpha, \beta, \gamma)$, will approach $f(x, y, z)$ uniformly as α, β and γ approach zero.

3. **The summability of the triple Fourier's series at points of continuity of the function developed.** Before proving the principal theorem in this connection, it will be necessary to prove several lemmas.

LEMMA 4. *Let R be a region in space, lying within the cube whose sides are $\alpha = \pm(\pi - \varrho_1)$, $\beta = \pm(\pi - \varrho_1)$, $\gamma = \pm(\pi - \varrho_1)$, where ϱ_1 is a small positive quantity, and such that no point of R lies within the sphere whose center is at 0 and whose radius is ϱ_2, where ϱ_2 is also a small positive quantity. Then if $\varphi(\alpha, \beta, \gamma)$ is a function that is finite and integrable* in the region R, the limit*

$$(44) \qquad \lim_{l, m, n \to \infty} \left[\frac{1}{l\,m\,n} \int\!\!\int\!\!\int_R \varphi(\alpha, \beta, \gamma) \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma}\, d\alpha\, d\beta\, d\gamma \right]$$

will exist and be equal to zero.

* Here and elsewhere throughout this section, when a function is said to be finite and integrable, it is meant that the function has an integral according to the definition of Lebesgue.

Represent by M the upper limit of the absolute value of R, and by ϱ the smaller of the two quantities ϱ_1 and $\varrho_2/\sqrt{3}$. Then, if ϵ is an arbitrarily small positive quantity, we can take a positive integer q, such that

$$\frac{8\pi^3 M}{q\sin^2\varrho} < \frac{\epsilon}{3}. \tag{45}$$

We shall show that for values of l, m and n greater than q, the expression in brackets in (44) is less than ϵ, and our lemma will be proved.

Divide the region R into two parts, R_1 and R_1', such that R_1 contains all the points for which $\alpha^2+\beta^2 < 2\varrho_2^2/3$, and R_1' all points for which $\alpha^2+\beta^2 \geqq 2\varrho_2^2/3$. If there are no points for which the first of these two inequalities holds, the region R_1' will coincide with R.

Since no point in the region R lies within the sphere of radius ϱ_2, whose center is at the origin, we have $\alpha^2+\beta^2+\gamma^2 \geqq \varrho_2^2$, and for points in R_1, $\alpha^2+\beta^2 < 2\varrho_2^2/3$; it follows on subtracting the second inequality from the first that $\gamma^2 > \varrho_2^2/3 > \varrho^2$, and hence $|\gamma| \geqq \varrho$ in R_1.

Dividing R_1' into two regions, R_2 and R_3, such that R_2 contains all points for which $|\alpha| < \varrho_2/\sqrt{3}$, and R_3 all points for which $|\alpha| \geqq \varrho_2/\sqrt{3}$, and proceeding as before, we find that, for points in R_2, $|\beta| \geqq \varrho$, and, for points in R_3, $|\alpha| \geqq \varrho$.

We then have

$$\begin{aligned}
&\left|\frac{1}{lmn}\int\int\int_{R_1} \varphi(\alpha,\beta,\gamma)\frac{\sin^2 l\alpha}{\sin^2\alpha}\frac{\sin^2 m\beta}{\sin^2\beta}\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\,d\beta\,d\gamma\right| \\
&\qquad < \frac{1}{lmn}\int\int\int_{R_1} \left|\varphi(\alpha,\beta,\gamma)\frac{\sin^2 l\alpha}{\sin^2\alpha}\frac{\sin^2 m\beta}{\sin^2\beta}\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\,d\beta\,d\gamma\right| \\
&\qquad < \frac{M}{lmn\sin^2\varrho}\int\int\int_{R_1} \frac{\sin^2 l\alpha}{\sin^2\alpha}\frac{\sin^2 m\beta}{\sin^2\beta}\,d\alpha\,d\beta\,d\gamma.
\end{aligned} \tag{46}$$

From Fejér's theorem* that

$$\frac{\pi}{2} = \frac{1}{n}\int_0^{\pi/2} \frac{\sin^2 n\alpha}{\sin^2\alpha}\,d\alpha,$$

together with (45) and (46), we have

$$\begin{aligned}
&\left|\frac{1}{lmn}\int\int\int_{R_1} \varphi(\alpha,\beta,\gamma)\frac{\sin^2 l\alpha}{\sin^2\alpha}\frac{\sin^2 m\beta}{\sin^2\beta}\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\,d\beta\,d\gamma\right| \\
&\qquad < \frac{8\pi^3 M}{n\sin^2\varrho} < \frac{\epsilon}{3}, \qquad (n \geqq q).
\end{aligned} \tag{47}$$

* L. c. p. 55.

If there are no points such that $\alpha^2 + \beta^2 < 2\varrho_2^2/3$, the inequality (47) still holds.

By similar methods, we find that analogous equations hold for $R_2\,(m > q)$, and for $R_3\,(1 > q)$. Combining these three equations, we have

$$\left|\frac{1}{lmn}\int\int\int_{R} q\,(\alpha, \beta, \gamma)\,\frac{\sin^2 l\alpha}{\sin^2\alpha}\,\frac{\sin^2 m\beta}{\sin^2\beta}\,\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\, d\beta\, d\gamma\right| < \epsilon, \quad (l, m, n > q),$$

and, as pointed out before, the lemma is proved.

LEMMA 5. *If g, g_1, h, h_1, k, k_1, are positive numbers less than π, the limit*

$$\lim_{l,m,n\to\infty}\left|\frac{1}{lmn\pi^3}\int_{-g_1}^{g}\int_{-h_1}^{h}\int_{-k_1}^{k}\frac{\sin^2 l\alpha}{\sin^2\alpha}\,\frac{\sin^2 m\beta}{\sin^2\beta}\,\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\, d\beta\, d\gamma\right| \tag{48}$$

will exist and be equal to unity.

The expression in brackets in (48) may be written in the form

$$\begin{aligned}&\frac{1}{lmn\pi^3}\int_{-\pi/2}^{\pi/2}\int_{-\pi/2}^{\pi/2}\int_{-\pi/2}^{\pi/2}\frac{\sin^2 l\alpha}{\sin^2\alpha}\,\frac{\sin^2 m\beta}{\sin^2\beta}\,\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\, d\beta\, d\gamma\\ &\pm\frac{1}{lmn\pi^3}\int\int\int_{R}\frac{\sin^2 l\alpha}{\sin^2\alpha}\,\frac{\sin^2 m\beta}{\sin^2\beta}\,\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\, d\beta\, d\gamma,\end{aligned} \tag{49}$$

where R is the region that must be added to, or subtracted from, the cube whose sides are $\alpha = \pm\frac{\pi}{2}$, $\beta = \pm\frac{\pi}{2}$, $\gamma = \pm\frac{\pi}{2}$ to produce the parallelopiped whose sides are $\alpha = g$, $\alpha = -g_1$, $\beta = h$, $\beta = -h_1$, $\gamma = k$, $\gamma = -k_1$. Since this region R satisfies the requirements of the region in Lemma 4, it follows that the second term of (49) approaches zero as l, m and n become infinite. Hence we need only to evaluate the first term, which may be written in the form

$$\left\{\frac{1}{l\pi}\int_{-\pi/2}^{\pi/2}\frac{\sin^2 l\alpha}{\sin^2\alpha}\,d\alpha\right\}\left\{\frac{1}{m\pi}\int_{-\pi/2}^{\pi/2}\frac{\sin^2 m\beta}{\sin^2\beta}\,d\beta\right\}\left\{\frac{1}{n\pi}\int_{-\pi/2}^{\pi/2}\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\gamma\right\},$$

which we know from Fejér's work to be equal to unity when l, m and n are positive integers. Hence its limit as l, m and n become infinite is unity, and

since the limit of the second term of (49) is zero, it follows that the limit of the entire expression (49) as l, m and n become infinite is unity.

LEMMA 6. *Let R be a region in space, lying within the cube whose sides are $\alpha = \pm(\pi - \varrho_1)$, $\beta = \mp(\pi - \varrho_1)$, $\gamma = \pm(\pi - \varrho_1)$, where ϱ_1 is a small positive quantity, and such that the point $\alpha = 0$, $\beta = 0$, $\gamma = 0$ lies within or on the boundary of R. Then, if $\varphi(\alpha, \beta, \gamma)$ is a function that is finite and integrable in R, the limit*

$$(50) \qquad \lim_{l, m, n \to \infty} \left[\frac{1}{l\,m\,n} \int\int\int_R \varphi(\alpha, \beta, \gamma) \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma} d\alpha\, d\beta\, d\gamma \right]$$

will exist and be equal to zero, provided

$$(51) \qquad \lim_{\alpha, \beta, \gamma \to +0} [\varphi(\alpha, \beta, \gamma)] = 0.$$

In view of (51) we may choose a quantity ϱ_2, so small that

$$(52) \qquad |\varphi(\alpha, \beta, \gamma)| < \frac{\epsilon}{2\pi^3} \qquad (\alpha^2 + \beta^2 + \gamma^2 < \varrho_2^2),$$

where ϵ is an arbitrarily small positive quantity. Now divide R into two parts, R_1 and R_2, where R_1 is the sphere with its center at the origin and radius ϱ_2, or as much of it as is included in R, and R_2 is the remainder of R. Then, from Lemma 4, it follows that

$$\lim_{l, m, n \to \infty} \left[\frac{1}{l\,m\,n} \int\int\int_{R_2} \varphi(\alpha, \beta, \gamma) \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma} d\alpha\, d\beta\, d\gamma \right] = 0,$$

and hence we can find a q so large that

$$(53) \qquad \left| \frac{1}{l\,m\,n} \int\int\int_{R_2} \varphi(\alpha, \beta, \gamma) \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma} d\alpha\, d\beta\, d\gamma \right| < \frac{\epsilon}{2} \quad (l, m, n > q).$$

Using (52) and Fejér's theorem,* we have

$$(54) \qquad \left| \frac{1}{l\,m\,n} \int\int\int_{R_1} \varphi(\alpha, \beta, \gamma) \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma} d\alpha\, d\beta\, d\gamma \right|$$

$$< \frac{\epsilon}{2\pi^3} \left\{ \frac{1}{l\,m\,n} \int\int\int_{R_1} \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma} d\alpha\, d\beta\, d\gamma \right\} < \frac{\epsilon}{2}.$$

* See previous footnote.

Combining (53) and (54), we find that the absolute value of the quantity in brackets in (50) is less than ϵ, ($l, m, n \geq q$), and hence the limit (50) exists and is equal to zero.

We wish now to consider the summability of the development of a function of three variables, $f(x, y, z)$, in a triple Fourier's series, i. e., the summability of the triple series

$$(55) \quad \sum_{l=1, m=1, n=1}^{\infty, \infty, \infty} \frac{1}{2^{E(1/l)+E(1/m)+E(1/n)}\pi^3} \cdot \int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(x', y', z')\, P_{lmn}(x, y, z, x', y', z')\, dx'\, dy'\, dz',$$

where

$$(56) \quad P_{lmn}(x, y, z, x', y', z') = \cos[(l-1)(x'-x)]\cos[(m-1)(y'-y)]\cos[(n-1)(z'-z)],$$

$E(s)$ representing the largest integer contained in s.

THEOREM III. *If the function $f(x, y, z)$ is finite and integrable in the region*

$$(57) \quad (-\pi \leq x \leq \pi, \ -\pi \leq y \leq \pi, \ -\pi \leq z \leq \pi),$$

the development of the function in a triple Fourier's series will be summable (C1) to the value of the function at every interior point of the region (57) at which the function is continuous.

For the series (55) we have*

$$\frac{S^{(1)}_{lmn}(x, y, z)}{A^{(1)}_{lmn}} = \frac{S^{(1)}_{lmn}(x, y, z)}{lmn}$$

$$= \frac{1}{8\, lmn\, \pi^3}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi}\int_{-\pi}^{\pi} f(x', y', z')\left(\frac{\sin\frac{l(x'-x)}{2}}{\sin\frac{x'-x}{2}}\right)^2 \cdot \left(\frac{\sin\frac{m(y'-y)}{2}}{\sin\frac{y'-y}{2}}\right)^2\left(\frac{\sin\frac{n(z'-z)}{2}}{\sin\frac{z'-z}{2}}\right) dx'\, dy'\, dz'.$$

* The reductions involved in arriving at this value, are exactly analogous to those for the simple series. See Fejér, l. c., p. 54.

Making the transformation

$$(58) \qquad x' - x = 2\alpha, \qquad y' - y = 2\beta, \qquad z' - z = 2\gamma,$$

we have

$$(59) \quad \frac{S^{(1)}_{lmn}(x, y, z)}{lmn} = \frac{1}{lmn\pi^3} \int_{-\frac{\pi+x}{2}}^{\frac{\pi-x}{2}} \int_{-\frac{\pi+y}{2}}^{\frac{\pi-y}{2}} \int_{-\frac{\pi+z}{2}}^{\frac{\pi-z}{2}} f(x+2\alpha, y+2\beta, z+2\gamma)$$

$$\cdot \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma} \, d\alpha \, d\beta \, d\gamma.$$

If we can show that the right-hand side of (59) approaches $f(x, y, z)$ as a limit as l, m and n become infinite, at any interior point of the region (57) at which $f(x, y, z)$ is continuous, our theorem will be proved.

Let

$$(60) \qquad \varphi(\alpha, \beta, \gamma) = f(x+2\alpha, y+2\beta, z+2\gamma) - f(x, y, z).$$

Then the right-hand side of (59) may be written

$$(61) \quad \begin{aligned} &\frac{1}{lmn\pi^3} \int_{-\frac{\pi+x}{2}}^{\frac{\pi-x}{2}} \int_{-\frac{\pi+y}{2}}^{\frac{\pi-y}{2}} \int_{-\frac{\pi+z}{2}}^{\frac{\pi-z}{2}} \varphi(\alpha, \beta, \gamma) \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma} \, d\alpha \, d\beta \, d\gamma \\ &+ \frac{f(x, y, z)}{lmn\pi^3} \int_{-\frac{\pi+x}{2}}^{\frac{\pi-x}{2}} \int_{-\frac{\pi+y}{2}}^{\frac{\pi-y}{2}} \int_{-\frac{\pi+z}{2}}^{\frac{\pi-z}{2}} \frac{\sin^2 l\alpha}{\sin^2 \alpha} \frac{\sin^2 m\beta}{\sin^2 \beta} \frac{\sin^2 n\gamma}{\sin^2 \gamma} \, d\alpha \, d\beta \, d\gamma. \end{aligned}$$

From the definition (60), $\varphi(\alpha, \beta, \gamma)$ is finite and integrable in the region of integration of the integrals in (61), and, moreover,

$$\lim_{\alpha, \beta, \gamma \to 0} [\varphi(\alpha, \beta, \gamma)] = 0,$$

provided $f(x, y, z)$ is continuous at the point (x, y, z). Hence, from Lemma 6, the first term of (61) approaches zero as l, m and n become infinite, if (x, y, z) is an interior point of the region (57), and if $f(x, y, z)$ is continuous at that point.

From Lemma 5, it follows that the second term of (61) and therefore the entire expression (61), and hence also the right-hand side of (59) approaches $f(x, y, z)$ as a limit as l, m and n become infinite, and therefore the development of the function in the triple Fourier's series is summable ($C1$) to the value of the function.

COROLLARY. *If $f(x, y, z)$ is finite and integrable in the region (57), its Fourier's development will be uniformly summable to $f(x, y, z)$ throughout any region R' whose boundary is interior to the boundary of a region of continuity of $f(x, y, z)$.*

By making slight modifications in Lemmas 4 and 6 and Theorem III, this corollary is easily established.

THEOREM IV. *If $f(x, y, z)$ satisfies the conditions of Theorem III, then for its Fourier's development*

$$L < \frac{S^{(1)}_{lmn}(x, y, z)}{lmn} < M, \qquad \left(\begin{array}{c} l, m, n = 1, 2, 3, \ldots \\ -\pi \leqq x \leqq \pi, \\ -\pi \leqq y \leqq \pi, \\ -\pi \leqq z \leqq \pi, \end{array}\right),$$

where L and M are the lower and upper limits respectively of $f(x, y, z)$ in the region (57).

From equation (59) we have

$$\frac{L}{lmn\pi^3}\int_{-\frac{\pi+x}{2}}^{\frac{\pi-x}{2}}\int_{-\frac{\pi+y}{2}}^{\frac{\pi-y}{2}}\int_{-\frac{\pi+z}{2}}^{\frac{\pi-z}{2}} \frac{\sin^2 l\alpha}{\sin^2\alpha}\frac{\sin^2 m\beta}{\sin^2\beta}\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\,d\beta\,d\gamma < \frac{S^{(1)}_{lmn}(x, y, z)}{lmn}$$

(62)

$$< \frac{M}{lmn\pi^3}\int_{-\frac{\pi+x}{2}}^{\frac{\pi-x}{2}}\int_{-\frac{\pi+y}{2}}^{\frac{\pi-y}{2}}\int_{-\frac{\pi+z}{2}}^{\frac{\pi-z}{2}} \frac{\sin^2 l\alpha}{\sin^2\alpha}\frac{\sin^2 m\beta}{\sin^2\beta}\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\alpha\,d\beta\,d\gamma,$$

but the integral in (62) may be written

$$\left(\frac{1}{l\pi}\int_{-\frac{\pi+x}{2}}^{\frac{\pi-x}{2}}\frac{\sin^2 l\alpha}{\sin^2\alpha}\,d\alpha\right)\left(\frac{1}{m\pi}\int_{-\frac{\pi+y}{2}}^{\frac{\pi-y}{2}}\frac{\sin^2 m\beta}{\sin^2\beta}\,d\beta\right)\left(\frac{1}{n\pi}\int_{-\frac{\pi+z}{2}}^{\frac{\pi-z}{2}}\frac{\sin^2 n\gamma}{\sin^2\gamma}\,d\gamma\right),$$

which is equal to unity,* and therefore the theorem follows at once.

* Fejér, l. c., p. 60.

4. **Application.** We shall now apply our results to a problem in the flow of heat. We wish to determine at any instant the temperature of any point of a rectangular parallelopiped whose initial temperature in known, and whose surface is maintained at the temperature zero. Let a be the length, b the width and c the height of the parallelopiped, and $f(x, y, z)$ a function giving the initial temperature of the parallelopiped at any point, when we take the origin at the lower, forward, left-hand corner of the parallelopiped, and let the x, y and z axes fall on the sides whose lengths are a, b and c respectively. The formal method of building up a solution for this type of problem gives us for the temperature of any point at any time*

$$(63)\qquad \sum_{l=1,m=1,n=1}^{\infty,\infty,\infty} u_{lmn}(x, y, z, t) = \frac{8}{abc} \sum_{l=1,m=1,n=1}^{\infty,\infty,\infty} a_{lmn}\, e^{-k\pi^2\left(\frac{l^2}{a^2}+\frac{m^2}{b^2}+\frac{n^2}{c^2}\right)t} \sin\frac{l\pi x}{a} \sin\frac{m\pi y}{b} \sin\frac{n\pi z}{c},$$

where k is an essentially positive quantity, and where

$$(64)\qquad a_{lmn} = \int_0^a\int_0^b\int_0^c f(x', y', z') \sin\frac{l\pi x'}{a} \sin\frac{m\pi y'}{b} \sin\frac{n\pi z'}{c}\, dx'\, dy'\, dz'.$$

In order to show that the expression (63) really furnishes a solution of the problem we must show that:

(1) the expression (63) converges and defines a continuous function of x, y, z and t, say $u(x, y, z, t)$, in the region

$$(65)\qquad 0 \leqq x \leqq a, \quad 0 \leqq y \leqq b, \quad 0 \leqq z \leqq c, \quad t > 0;$$

(2) the function $u(x, y, z, t)$ satisfies the equation

$$(66)\qquad \frac{\partial u}{\partial t} = k\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right)$$

* Carslaw, *Introduction to the mathematical theory of the conductivity of heat in solids*, p. 108.

throughout the region (65); and that

$$(3)\qquad \lim_{\substack{t\to+0\\ x\to x_1,\, y\to y_1\, z\to z_1}} u(x,y,z,t) = f(x_1,y_1,z_1),$$

where (x_1, y_1, z_1) is a point within a region throughout which $f(x, y, z)$ is continuous; and that as t approaches zero through positive values, $u(x,y,z,t)$ remains finite as x, y and z approach the coördinates of any point in the region

$$(67)\qquad 0 \leqq x \leqq a, \qquad 0 \leqq y \leqq b, \qquad 0 \leqq z \leqq c.$$

We shall first show that the series (63) is convergent in the region (65). Due to the presence of the convergence factor $e^{-k\pi^2(l^2/a^2+m^2/b^2+n^2/c^2)t}$ in (63), we have for the general term of the series

$$(68)\qquad |u_{lmn}(x,y,z,t)| < \frac{K}{l^2 m^2 n^2}$$

in the region

$$(69)\qquad 0 \leqq x \leqq a, \qquad 0 \leqq y \leqq b, \qquad 0 \leqq z \leqq c, \qquad 0 < t_0 < t,$$

where K is a positive constant and t_0 an arbitrarily small positive constant, provided $f(x, y, z)$ is finite and integrable in (67). The right-hand side of (68) is the general term of a convergent triple series of positive terms, and hence the series (63), of which the left-hand side of (68) is the general term, is absolutely convergent in the region (69), and, since t_0 is arbitrary, in the region (65).

The terms of (63) are functions of x, y and z in the region (69) and hence the series (63) is uniformly convergent throughout the region (69) by virtue of Weierstrass' test, extended for the triple series. Since the terms of the series are continuous throughout the region (69), it follows that the series defines a continuous function throughout that region, and, since t_0 is arbitrary, throughout the region (65). Hence the condition (1) is satisfied.

We shall now show that the condition (2) is satisfied, i.e., that the series (63) satisfies the equation (66). Each term obviously satisfies the equation, and therefore the series will satisfy it provided we have a right to form the derivatives involved in differentiating (63) by differentiating term by term. We may form the derivative of a triple series convergent in a certain region by differentiating term by term if the derived series is uniformly convergent throughout the region and defines a continuous function there. We have shown that the original series converges in the region (65), and it is easily

shown that the derived series is uniformly convergent and continuous in this region by a method analogous to that used in proving the uniform convergence and continuity of the original series in that region. Hence condition (2) is satisfied.

In order to show that condition (3) is satisfied it will be necessary to make use of Theorem II on convergence factors. The convergence factors which occur in the series (63) are functions of one variable only, whereas those involved in Theorem II were functions of three variables, for particular values of l, m and n. We may put the convergence factors of (63) in a form where they are functions of three variables by making the transformations

$$\alpha = \frac{k\pi^2 t}{a^2}, \qquad \beta = \frac{k\pi^2 t}{b^2}, \qquad \gamma = \frac{k\pi^2 t}{c^2}.$$

The convergence factors then have the form

$$e^{-(l^2\alpha + m^2\beta + n^2\gamma)}. \tag{70}$$

We shall first prove the first part of condition (3), i. e., that, as t approaches zero through positive values and as x, y and z approach x_1, y_1, z_1, where (x_1, y_1, z_1) is a point of continuity of the function, the function $u(x, y, z, t)$ approaches $f(x_1, y_1, z_1)$. In order to prove this we shall have to show that the series (63), without the convergence factors, satisfies the conditions of Theorem II, and that the convergence factors (70) satisfy conditions (a)—(d) of Lemma 3 and (e) and (f) of Theorem II. It will then follow from the corollary to Theorem II that $u(x, y, z, t)$ has the desired property.

It follows from the corollary to Theorem III, by a change of variable, that the series (63) without the convergence factors is uniformly summable throughout a region whose boundary is interior to that of a region of continuity of $f(x, y, z)$, and the condition on the series is satisfied.

Hence it remains only to show that the convergence factors (70) satisfy the conditions (a)—(i) inclusive of Theorem II. It is obvious that conditions (d), (e) and (f) are satisfied, and hence only conditions (a), (b), (c), (g), (h) and (i) remain.

Consider first condition (c). Since

$$\left| e^{-(l^2\alpha_0 + m^2\beta_0 + n^2\gamma_0)} \right| < \frac{K}{l^4 m^4 n^4}, \qquad (l, m, n = 1, 2, 3, \ldots), \tag{71}$$

where α_0, β_0, γ_0 and K are positive constants, we have

$$\sum_{k=1}^{\infty} k \left| e^{-(l^2\alpha_0 + m^2\beta_0 + k^2\gamma_0)} \right| < \frac{K}{l^4 m^4} \sum_{k=1}^{\infty} \frac{1}{k^3} = \frac{K_1}{l^4 m^4},$$

where K_1 is a positive constant. Hence

$$\lim_{l, m \to \infty} l\, m \sum_{k=1}^{\infty} k \left| e^{-(l^2\alpha + m^2\beta + k^2\gamma)} \right| = 0, \qquad (\alpha, \beta, \gamma > 0).$$

The proof for condition (b) is carried out in a manner analogous to the above proof for condition (c).

Consider now condition (a). We have

$$\sum_{\lambda_1, \mu_1, \nu_1}^{\lambda_2, \mu_2, \nu_2} i\, j\, k \left| \triangle^{\frac{2}{2}}_{\frac{2}{2}}\, e^{-(i^2\alpha + j^2\beta + k^2\gamma)} \right| = \left[\sum_{i=\lambda_1}^{\lambda_2} i \left(e^{-i^2\alpha} - 2\, e^{-(i+1)^2\alpha} + e^{-(i+2)^2\alpha} \right) \right]$$

$$\cdot \left[\sum_{j=\mu_1}^{\mu_2} j \left(e^{-j^2\beta} - 2\, e^{-(j+1)^2\beta} + e^{-(j+2)^2\beta} \right) \right] \cdot \left[\sum_{k=\nu_1}^{\nu_2} k \left(e^{-k^2\gamma} - 2\, e^{-(k+1)^2\gamma} + e^{-(k+2)^2\gamma} \right) \right],$$

and therefore condition (a) will hold for some K such that

$$(72) \qquad \sum_{i=1}^{\infty} i \left(e^{-i^2 u} - 2\, e^{-(i+1)^2 u} + e^{-(i+2)^2 u} \right) < \sqrt[3]{K}, \qquad (u > 0).$$

From the law of the mean it follows that

$$(73) \quad e^{-i^2 u} - 2\, e^{-(i+1)^2 u} + e^{-(i+2)^2 u} = e^{-(i+\theta)^2 u} \{ 4 (i + \theta)^2 u^2 - 2 u \}, \quad (0 < \theta < 2).$$

From this we see that the terms of the series on the left-hand side of (72) are negative for all positive integral values of i when $i + 2 > 1/\sqrt{2u}$, and positive for all positive integral values of i when $i > 1/\sqrt{2u}$. Therefore the series consists of a group of negative terms followed by two terms whose signs may be plus or minus, followed by a group of positive terms. It is readily seen from (73) that each term of the series (72) remains less than some positive constant M, and hence the inequality (72) holds for some K provided any sequence we may choose, consisting entirely of positive or of negative terms of the series in (72), remains less in absolute value than some positive constant.

Consider such a sequence, giving i all integral values from p to q, where p and q are positive integers. The sum of this sequence is

$$(74) \qquad p\, e^{-p^2 u} - (p - 1)\, e^{-(p+1)^2 u} - (q + 1)\, e^{-(q+1)^2 u} + q\, e^{-(q+2)^2 u},$$

which differs from

$$p\,[e^{-p^2 u} - e^{-(p+1)^2 u}] - (q+1)\,[e^{-(q+1)^2 u} - e^{-(q+2)^2 u}] \tag{75}$$

by $e^{-(p+1)^2 u} - e^{-(q+2)^2 u}$, which can never exceed unity. But by the law of the mean,

$$\left| m\,(e^{-m^2 u} - e^{-(m+1)^2 u}) \right| < 2\,(m+\theta)^2 u\, e^{-(m+\theta)^2 u}, \quad (0 < \theta < 1), \tag{76}$$

the right-hand side of which always remains less than some positive constant for all positive integral values of m and all positive values of u. Therefore the left-hand side of (76), and therefore also (75) and (74), remains less than some positive constant for all values of $u > 0$; hence, as pointed out before, (72) holds, and hence the condition (a) is satisfied.

It is obvious from the fact that expression (74) is finite and from condition (f) that conditions (g) and (h) are satisfied. That condition (i) is satisfied follows from condition (a) and the equation (71). Hence the convergence factors satisfy the conditions of Theorem II, and as pointed out before the first part of condition (c) of this article is satisfied.

If now remains only to prove that, as t approaches zero through positive values and x, y and z approach the coordinates of any point in the region (67), the function $u\,(x, y, z, t)$ remains finite. Since the convergence factors satisfy the conditions of Lemma 3, it follows from that lemma that, for all positive values of α, β and γ or, what is the same thing, for all values of $t > 0$,

$$\sum_{i=1,\,j=1,\,k=1}^{\infty,\,\infty,\,\infty} S^{(1)}_{ijk}\,(x, y, z)\, \triangle^{2}_{\substack{2\\2}}\, e^{-(i^2\alpha + j^2\beta + k^2\gamma)} \tag{77}$$

converges to the same value as the series (63), and hence, if we prove that (77) remains fimite for all values of $t > 0$ when x, y and z approach the coördinates of any point in the region (67), we shall have proved that the series (63) satisfies this requirement also.

Since $f(x, y, z)$ is finite and integrable in (67), we have from Theorem IV

$$\left| S^{(1)}_{ijk}\,(x, y, z) \right| < C\,i\,j\,k \qquad \begin{pmatrix} 0 \leq x \leqq a,\ 0 \leqq y \leqq b,\ 0 \leq z \leq c \\ i, j, k = 1, 2, 3, \ldots \end{pmatrix},$$

where C is a positive constant. Making use of this inequality and the fact that the convergence factors (70) satisfy condition (a) of Lemma 3, we have

$$\left| \sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} S_{ijk}^{(1)}(x, y, z) \triangle_{2}^{\substack{2\\2}} e^{-(i^2\alpha + j^2\beta + k^2\gamma)} \right|$$

$$< \sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} \left| S_{ijk}^{(1)} \right| \cdot \left| \triangle_{2}^{\substack{2\\2}} e^{-(i^2\alpha + j^2\beta + k^2\gamma)} \right|$$

$$< \sum_{i=1, j=1, k=1}^{\infty, \infty, \infty} C\, i\, j\, k \left| \triangle_{2}^{\substack{2\\2}} e^{-(i^2\alpha + j^2\beta + k^2\gamma)} \right| < CK,$$

$$\left(\substack{\alpha,\ \beta,\ \gamma > 0 \\ 0 < x < a,\ 0 \leq y < b,\ 0 < z \leq c} \right).$$

and hence the final condition is satisfied and the series (63) furnishes a solution of the physical problem.

CINCINNATI, OHIO,
December 1921.

ON THE MINIMIZING OF A CLASS OF DEFINITE INTEGRALS.*

By Paul R. Rider.

Introduction. Several special problems in the calculus of variations lead to the consideration of a definite integral of the form

$$\int_{x_0}^{x_1} \frac{[1+y'^2(x)]^m}{y''(x)}\, dx. \tag{1}$$

For example, Euler's historic problem of finding the curve which with its evolute encloses a minimum area† gives rise to the particular case $m = 2$ in the above integral. The case $m = 1$ arises in obtaining the curve which with its caustic encloses a minimum area.‡ In finding the curves of minimum mean radius of curvature with respect to the arc and with respect to the abscissa, we are led to the cases $m = 2$ and $m = \frac{3}{2}$ respectively. It would thus seem desirable to develop a theory for the minimizing of the integral (1), to which, moreover, peculiar interest attaches on account of the fact that the second derivative appears in the integrand, comparatively few problems of that kind having been completely solved.

It is the purpose of this paper to derive such a theory. In section 1 are obtained the equations of the extremals, or curves that minimize the integral (1). In section 2 it is shown that these curves actually do furnish a minimum value for the integral. The question of the determination of the arbitrary constants that occur in the solution of section 1 is studied in section 3. Sections 4—6 treat various cases of variable end conditions. Finally, in section 7, a concrete illustration is given by means of the curve of minimum mean radius of curvature with respect to x.

1. **The minimizing curves.** Our problem is that of finding a curve $y = y(x)$ which joins the point $P_0\,(x_0, y_0)$ to the point $P_1\,(x_1, y_1)$ $(x_0 < x_1)$, and which gives to the integral (1) a smaller value than any neighboring curve. Let us assume that such a curve exists, and consider comparison curves of

* Presented to the American Mathematical Society, April 14, 1922.

† See Todhunter, Researches in the calculus of variations, chapter XIII.

‡ See Rider, The minimum area between a curve and its caustic, Bull. Amer. Math. Soc., vol. 27, p. 279; also Dunkel, The curve which with its caustic encloses the minimum area, Washington Univ. Studies, vol. 8, scientific series, p. 183.

the form $y = y(x) + \epsilon\eta(x)$. The value of the integral (1) for one of these comparison curves is

$$I(\epsilon) = \int_{x_0}^{x_1} \frac{[1 + (y' + \epsilon\eta')^2]^m}{y'' + \epsilon\eta''}\, dx.$$

Since the comparison curves reduce to the minimizing curves for $\epsilon = 0$, we must have $I'(0) = 0$ for a minimum. We find that

$$(2) \qquad I'(0) = \int_{x_0}^{x_1} \left[\frac{2my'(1 + y'^2)^{m-1}}{y''}\eta' - \frac{(1 + y'^2)^m}{y''^2}\eta'' \right] dx,$$

which, by the usual integration by parts employed in the calculus of variations, becomes

$$(3) \quad I'(0) = \left[\frac{2my'(1 + y'^2)^{m-1}}{y''} + \frac{d}{dx}\frac{(1 + y'^2)^m}{y''^2} \right] \eta \Bigg|_{x_0}^{x_1} - \frac{(1 + y'^2)^m}{y''^2}\eta' \Bigg|_{x_0}^{x_1}$$

$$- \int_{x_0}^{x_1} \left[\frac{d}{dx}\frac{2my'(1 + y'^2)^{m-1}}{y''} + \frac{d^2}{dx^2}\frac{(1 + y'^2)^m}{y''^2} \right] \eta\, dx.$$

If it is prescribed that the end points P_0 and P_1 are fixed and that the minimizing curve shall have the slopes p_0 and p_1 at P_0 and P_1 respectively, then we choose as comparison curves those for which η and η' are zero at the end points, and therefore the first two expressions on the right side of equation (3) vanish. If the integral in (3) is also to vanish, it follows from the fundamental lemma of the calculus of variations that the coefficient of η in the integrand must be equal to zero. This gives a differential equation whose first integral is

$$\frac{2mp(1 + p^2)^{m-1}}{q} + \frac{d}{dx}\frac{(1 + p^2)^m}{q^2} = \text{constant} = 2A,$$

where we have set $p = y'$, $q = y''$. This reduces to

$$\frac{2mp}{1 + p^2} - \frac{q'}{q^2} = \frac{Aq}{(1 + p^2)^m}.$$

Multiplying by $dp = q\,dx$, we get

$$\frac{2\,mp\,dp}{1+p^2} - \frac{dq}{q} = \frac{Aq\,dp}{(1+p^2)^m},$$

or

$$d\log f = \frac{A\,dp}{f}, \tag{4}$$

where

$$f = \frac{(1+p^2)^m}{q}.$$

Integrating (4), replacing f by $(1+p^2)^m/q$, and solving for q, we obtain

$$q = \frac{(1+p^2)^m}{Ap+B}. \tag{5}$$

Since $q = dp/dx$, and $dy = p\,dx$, we find that

$$\left\{\begin{aligned} x &= \int_0^p \frac{Ap+B}{(1+p^2)^m}\,dp + C, \\ y &= \int_0^p \frac{Ap^2+Bp}{(1+p^2)^m}\,dp + D, \end{aligned}\right. \tag{6}$$

in which C and D are arbitrary constants. Thus

The extremals are given by the parametric equations (6), *the parameter p being the slope of the curve at the point* (x, y).

It is not difficult to show that changing the sign of A merely reflects a given curve in the line $x = C$, and that changing the sign of B reflects it in the line $y = D$. Consequently, if $A = 0$ the curve is symmetric with respect to the line $x = C$, and if $B = 0$ it is symmetric with respect to the line $y = D$.

As a typical case we select $A < 0 < B$. Suppose $B = -rA$, $r > 0$. Then q is positive and the curve is concave upward as p increases from $-\infty$ to r, for which value q becomes infinite and the curve has a cusp. For $r < p < \infty$, q is negative and the curve is concave downward. By a study of dx/dp we discover that as p increases from $-\infty$ to r, x increases up to a maximum value at the cusp, and then recedes as p increases from r to ∞. By a study of dy/dp we learn that as p increases from $-\infty$ to 0, y decreases to a minimum value, and then as p increases to r, y increases to a maximum value at the cusp, and then decreases as p increases from r to ∞.

2. **Sufficient conditions.** To show that under certain slight restrictions the curves (6) actually do give a minimum value to the integral (1), let us consider the difference between the value of the integral for a comparison curve $y = y(x) + \omega(x)$ and its value for the extremal. If we denote this difference by ΔI, it follows by means of Taylor's theorem, that

$$\Delta I = \int_{x_0}^{x_1} \left\{ \frac{2my'(1+y'^2)^{m-1}}{y''}\omega' - \frac{(1+y'^2)^m}{y''^2}\omega'' \right\} dx$$

$$+ \int_{x_0}^{x_1} \left\{ \left[\frac{2m(m-1)\bar{y}'^2(1+\bar{y}'^2)^{m-2}}{\bar{y}''} + \frac{m(1+\bar{y}'^2)^{m-1}}{\bar{y}''} \right] \omega'^2 \right.$$

$$\left. - \frac{2m\bar{y}'(1+\bar{y}'^2)^{m-1}}{\bar{y}''^2}\omega'\omega'' + \frac{(1+\bar{y}'^2)^m}{\bar{y}''^3}\omega''^2 \right\} dx,$$

where $\bar{y}' = y' + \theta\omega'$, $\bar{y}'' = y'' + \theta\omega''$ $(0 < \theta < 1)$.

The first integral, however, vanishes, since the functions y' and y'' refer to the extremals. (Compare equation (2).) Thus, after a simple reduction of the second integral, we have

$$\Delta I = \int_{x_0}^{x_1} \frac{(1+\bar{y}'^2)^m}{\bar{y}''} \left\{ \left[\frac{m\bar{y}'\omega'}{1+\bar{y}'^2} - \frac{\omega''}{\bar{y}''} \right]^2 + \left[\frac{(m-1)\bar{y}'\omega'}{1+\bar{y}'^2} \right]^2 \right.$$

$$\left. + \frac{\omega'^2}{(1+\bar{y}'^2)^2} + \frac{(m-1)\omega'^2}{1+\bar{y}'^2} \right\} dx.$$

If we assume that our extremals are concave upward, that is $y'' > 0$, and consider only comparison curves which are concave upward, we see that ΔI is positive for $m \geq 1$. If $m = 0$, the sum of the last three terms in the integrand is zero; consequently ΔI is also positive if $m = 0$. Therefore:

If $m \geq 1$ or $m = 0$ the curves (6) will give a smaller value to the integral (1) than any neighboring curves for which y'' is positive.

We can also reduce ΔI to the form

$$\Delta I = \frac{1}{2} \int_{x_0}^{x_1} \frac{(1+\bar{y}'^2)^m}{\bar{y}''} \left\{ \left[\frac{2m\bar{y}'\omega'}{1+\bar{y}'^2} - \frac{\omega''}{\bar{y}''} \right]^2 + \frac{2m(1-\bar{y}'^2)}{(1+\bar{y}'^2)^2}\omega'^2 + \frac{\omega''^2}{\bar{y}''^2} \right\} dx.$$

This is positive if $m(1-\bar{y}'^2) \geq 0$. But $m(1-\bar{y}'^2)$ will be positive if $m > 0$ and $|\bar{y}'| < 1$, or if $m < 0$ and $|\bar{y}'| > 1$. Thus we can add to the above conclusion the following:

If $0 < m < 1$, and we consider only curves having slopes not greater than one in absolute value, the portions of the curves (6) *for which* $|p| \leqq 1$ *will give a minimum value to the integral* (1).

If $m < 0$, and we consider only curves having slopes not less than one in absolute value, the portions of the curves (6) *for which* $|p| \geqq 1$ *will give a minimum value to the integral* (1).

3. **The arbitrary constants and the minimum value of the integral.** If we make use of the conditions that the minimizing curve passes through the point P_0 with the slope p_0 and through the point P_1 with the slope p_1, we obtain from (6) the following four linear equations to solve for the arbitrary constants A, B, C, D:

$$(7) \qquad \begin{cases} \beta_0 A + \alpha_0 B + C = x_0, \\ \gamma_0 A + \beta_0 B + D = y_0, \\ \beta_1 A + \alpha_1 B + C = x_1, \\ \gamma_1 A + \beta_1 B + D = y_1, \end{cases}$$

where

$$(8) \qquad \alpha_i = \int_0^{p_i} \frac{dp}{(1+p^2)^m}, \qquad \beta^i = \int_0^{p_i} \frac{p\,dp}{(1+p^2)^m}, \qquad \gamma_i = \int_0^{p_i} \frac{p^2\,dp}{(1+p^2)^m}.$$

The determinant of the coefficients of equations (7) has the value

$$\Delta = \left[\int_{p_0}^{p_1} \frac{p\,dp}{(1+p^2)^m}\right]^2 - \int_{p_0}^{p_1} \frac{dp}{(1+p^2)^m} \cdot \int_{p_0}^{p_1} \frac{p^2\,dp}{(1+p^2)^m},$$

which, by Schwarz's inequality, is negative. Therefore, *we can always solve for the arbitrary constants A, B, C, D, and they are uniquely determined in terms of $x_0, y_0, p_0, x_1, y_1, p_1$.*

Since we wish to use the values of A and B in the next section, we solve for these values here, obtaining

$$(9) \qquad \begin{cases} A = \dfrac{1}{\Delta}\left[(x_1 - x_0)\displaystyle\int_{p_0}^{p_1} \frac{p\,dp}{(1+p^2)^m} - (y_1 - y_0)\int_{p_0}^{p_1} \frac{dp}{(1+p^2)^m}\right], \\ B = \dfrac{1}{\Delta}\left[(y_1 - y_0)\displaystyle\int_{p_0}^{p_1} \frac{p\,dp}{(1+p^2)^m} - (x_1 - x_0)\int_{p^0}^{p_1} \frac{p^2\,dp}{(1+p^2)^m}\right]. \end{cases}$$

By using equations (6) in (1), we discover that the minimum value of the integral I is

(10) $$I = A(y_1 - y_0) + B(x_1 - x_0).$$

4. **Variable slope at an end point.** Let us now consider the case in which the slope at one end point, say P_1, is not fixed. In such a case the value of η' in equation (3) is not zero for $x = x_1$, and therefore, if $I'(0)$ is to vanish, *we must have*

$$\frac{(1+y'^2)^m}{y''^2} = 0$$

at P_1: that is, *there must be a cusp at this point.* We readily find from (5) that *the curve must have at the point* P_1 *the slope* $-B/A$.

5. **End point variable on a vertical line.** Let us now suppose that the slope of the curve at the point P_1 has a fixed finite value, but that the point itself is allowed to move along the line $x = x_1$. To find the conditions that will give the minimum value to the integral I we set $\partial I/\partial y_1 = 0$. Thus from (10) and (9) we get

$$\frac{\partial I}{\partial y_1} = A + (y_1 - y_0)\frac{\partial A}{\partial y_1} + (x_1 - x_0)\frac{\partial B}{\partial y_1} = \frac{2A}{\Delta} = 0.$$

Since Δ can not be infinite if p_0 and p_1 are finite, A *must be equal to zero.*

From the first equation of (9) we find that the value of y_1 which will make $A = 0$ is

$$y_1 = y_0 + \frac{\beta_1 - \beta_0}{\alpha_1 - \alpha_0}(x_1 - x_0),$$

where the α's and β's are defined by (8).

That this value of y_1 does furnish a minimum value for the integral I in the case under consideration is seen by observing that the second derivative,

$$\frac{\partial^2 I}{\partial y_1^2} = -\frac{2}{\Delta}\int_{p_0}^{p_1}\frac{dp}{(1+p^2)^m},$$

is positive, since $\Delta < 0$.

When $A = 0$, we obtain from (7) or (9), $B = (x_1 - x_0)/(\alpha_1 - \alpha_0)$, and thus I reduces to

$$I = B(x_1 - x_0) = \frac{(x_1 - x_0)^2}{\displaystyle\int_{p_0}^{p_1}\frac{dp}{(1+p^2)^m}}.$$

The fact that A must be zero if the end point is variable on a vertical line can also be shown as follows:

The expression for I can easily be changed to the form

$$I = \int_{p_0}^{p_1} (1+p^2)^m \left(\frac{dx}{dp}\right)^2 dp.$$

In this integral the limits are constant, since the curve has a fixed slope p_0 at the point P_0 and crosses the line $x = x_1$ with a fixed slope p_1. Moreover, the values of x corresponding to p_0 and p_1 are constant. Therefore we can obtain, by the usual process of the calculus of variations, for the differential equation to be satisfied by the minimizing curve,

$$(1+p^2)^m \frac{dx}{dp} = \text{constant}.$$

If we designate this constant by B and integrate to obtain x, we get

$$x = \int_0^p \frac{B\,dp}{(1+p^2)^m} + C.$$

Since $dy = p\,dx$, the value of y is the same as in (6) with A set equal to zero.

6. **End point variable on a horizontal line.** If the point P_1 is allowed to vary along the line $y = y_1$, but intersects that line with a fixed slope p_1, we proceed as in the preceding section. We find that

$$\frac{\partial I}{\partial x_1} = (y_1 - y_0)\frac{\partial A}{\partial x_1} + (x_1 - x_0)\frac{\partial B}{\partial x_1} + B = \frac{2B}{\Delta} = 0.$$

and therefore *B must vanish.*

From (9) we find that the value of x_1 which will cause B to vanish is

$$x_1 = x_0 + \frac{\beta_1 - \beta_0}{\gamma_1 - \gamma_0}(y_1 - y_0).$$

That this value of x_1 does give a minimum value to I is shown by the fact that

$$\frac{\partial^2 I}{\partial x_1^2} = -\frac{2}{\Delta}\int_{p_0}^{p_1} \frac{p^2\,dp}{(1+p^2)^m} > 0.$$

If $B = 0$ we see that $A = (y_1 - y_0)/(x_1 - x_0)$, and that

$$I = A(y_1 - y_0) = \frac{(y_1 - y_0)^2}{\displaystyle\int_{p_0}^{p_1} \frac{p^2\,dp}{(1+p^2)^m}}.$$

7. The curve of minimum mean radius of curvature. As an illustration, let us examine the curve whose mean radius of curvature with respect to x is a minimum. If we set up the integral which expresses the mean radius of curvature, we obtain, as was noted in the introduction, the special case $m = \dfrac{3}{2}$ in (1).

It follows from (6) that the equations of the extremals are

$$x = \int \frac{Ap + B}{(1+p^2)^{\frac{3}{2}}}\,dp, \qquad y = \int \frac{Ap^2 + Bp}{(1+p^2)^{\frac{3}{2}}}\,dp.$$

To integrate, it is convenient to set $p = \tan \tau$. We get

$$(10) \qquad \begin{cases} x = -A\cos\tau + B\sin\tau + C, \\ y = A\log(\sec\tau + \tan\tau) - A\sin\tau - B\cos\tau + D. \end{cases}$$

If $A = 0$, equations (10) reduce to $x - C = B\sin\tau$, $y - D = -B\cos\tau$, a circle.

WASHINGTON UNIVERSITY,
ST. LOUIS, MO.

A PYTHAGOREAN FUNCTIONAL EQUATION.*

BY EINAR HILLE.

1. Introduction. It is well known that

$$|\sin(x+iy)|^2 = \sin^2 x + \sinh^2 y. \tag{1}$$

This relation can also be written

$$|\sin(x+iy)|^2 = |\sin x|^2 + |\sin iy|^2, \tag{2}$$

or, in other words, the function $\sin z$ $(z = x + iy)$ satisfies the functional equation

$$|f(z)|^2 = |f(x)|^2 + |f(iy)|^2, \tag{3}$$

a *Pythagorean Theorem* in the theory of functions.

It is easy to see that $\sin z$ is not the only solution of (3). Evidently $C \sin z$ will also satisfy where C is an arbitrary complex constant. A special solution is found to be given by Cz. It is not hard to verify that

$$C\frac{\sin az}{a} \tag{4}$$

is a solution of (3) where C is an arbitrary constant and a is either real or purely imaginary. We shall prove that this is the general solution, if we restrict ourselves to analytic solutions.

In § 2 we reduce the problem to the solving of a differential equation in two independent and two dependent variables. This equation we integrate by a specialization of one of the independent variables which is equivalent to assuming the solutions of (3) single-valued and analytic in the neighbourhood of the origin and at that point.

In § 3 this assumption is justified in different ways. This section also contains an *a priori* discussion of the singularities of solutions of the functional equation.

* Presented to the American Mathematical Society, October 28, 1922. The results of this paper may be considered trivial, but the author believes that the methods used are of some interest and may find applications to more general problems.

2. The reduction. Let us introduce

$$(5)\qquad [f(z)]^2 = F(z);\qquad |F(z)| = R(z);\qquad |F(x)| = G(x);\qquad |F(iy)| = H(y),$$

which carries (3) over into the form

$$(6)\qquad R(z) = G(x) + H(y),$$

which is easy to solve. We can define $G(x)$ and $H(y)$ arbitrarily and $R(z)$ will be defined by (6). But when is the solution so obtained, the absolute value of an analytic function?

It is necessary that log $[R(z)]$ *is a harmonic function.* Hence

$$(7)\qquad \Delta\{\log[R(z)]\} = 0,$$

where Δ stands for the Laplacian. If we substitute for $R(z)$ its expression (6), form the Laplacian and simplify, we get

$$(8)\qquad [G(x)+H(y)]\,[G''(x)+H''(y)] - [G'(x)]^2 - [H'(y)]^2 = 0,$$

where the primes denote differentiation with respect to the argument of the function in question. Here we have a mixed differential equation containing two unknown functions $G(x)$ and $H(y)$. In order to obtain $G(x)$ we specialize y. The most natural specialization would certainly be to equate y to 0. This is allowable, at least if the solutions of (3) are single-valued and analytic at the origin.

We assume that this is the case and shall justify our assumption in the next section. Suppose a solution $f(z)$ of (3) is given by a power series

$$(9)\qquad f(z) = c_0 + c_1 z + c_2 z^2 + \cdots,$$

convergent for small values of z; $f(z)$ being analytic we can put $z = 0$ in equation (3) obtaining

$$(10)\qquad f(0) = c_0 = 0.$$

Further put $|c_1| = c$, then

$$(11)\qquad R(z) = c^2|z|^2\{1 + p(|z|^2)\},$$

where $p(|z|^2)$ is a power series in $|z|^2$ vanishing with z. Hence

$$H(0) = H'(0) = 0; \qquad H''(0) = 2c^2. \tag{12}$$

Substituting $y = 0$ in (8) we obtain

$$G(x)\,G''(x) - [G'(x)]^2 + 2c^2\,G(x) = 0. \tag{13}$$

If we put

$$G(x) = u^2, \tag{14}$$

the differential equation (13) is carried over into

$$uu'' - (u')^2 + c^2 = 0. \tag{15}$$

This equation is integrable by elementary methods and the general solution is found to be

$$u = c\,\frac{\sin a(x-\alpha)}{a}, \tag{16}$$

where a and α are arbitrary constants. The limiting value for $a \to 0$, namely $c(x-\alpha)$, is also a solution. We want to get the solution which vanishes at the origin. Hence $\alpha = 0$. Further $u^2 = G(x) > 0$ which shows that u must be a real function of x. c being real, we have to take a either real or purely imaginary. Thus

$$G(x) = |f(x)|^2 = c^2\,\frac{\sin^2 ax}{a^2}. \tag{17}$$

In a similar manner we show that

$$H(y) = |f(iy)|^2 = d^2\,\frac{\sin^2 by}{b^2}. \tag{18}$$

But $f(z)$ is analytic at the origin, thus

$$c = d, \qquad b = ia. \tag{19}$$

Hence

$$|f(z)|^2 = c^2\left[\frac{\sin^2 ax}{a^2} + \frac{\sinh^2 ay}{a^2}\right] = c^2\left|\frac{\sin az}{a}\right|^2. \tag{20}$$

If the absolute value of an analytic function is known in a simply connected region, then the argument is determined uniquely up to an additive real constant. Consequently

$$f(z) = C\frac{\sin az}{a}, \tag{21}$$

where C is an arbitrary complex constant, $|C| = c$, and a is a constant, either real or purely imaginary. When $a = 0$, formula (21) is understood to mean $f(z) = Cz$.

3. **The singularities.** We have assumed in § 2 that an arbitrary solution, $f(z)$, of (3) is single-valued and analytic at the origin. This assumption can be justified in different ways.

We can use equation (8) for this purpose. Suppose that every analytic solution of (3) is single-valued and analytic in the neighbourhood of some point on the axis of imaginaries. This point may, of course, differ from one solution to another. Suppose that for the solution $f(z)$ we may take $y = y_0$ and assume

$$H(y_0) = A, \quad H'(y_0) = B, \quad H''(y_0) = C.$$

Here $A > 0$, B and C are real. The equation (8) becomes

$$[G(x)+A][G''(x)+C]-[G'(x)]^2-B^2 = 0. \tag{22}$$

This equation is also integrable in terms of elementary functions. The general integral is of the form

$$G(x) = a\sin(bx+c)+d, \tag{23}$$

where two of the constants are expressible in terms of the other two and A, B, C. Their values are of no interest in this connection.

Supposing $f(z)$ single-valued and analytic even in the neighbourhood of $z = x_0$, we obtain in a similar manner

$$H(y) = \alpha\sin(\beta y+\gamma)+\delta \tag{24}$$

where, of course, α, β, γ, δ are not independent of a, b, c, d. Substitute the expressions (23) and (24) in equation (8) and we obtain

$$\begin{aligned} (25)\quad -a^2b^2-\alpha^2\beta^2-(d+\delta)\,[ab^2\sin(bx+c)+\alpha\beta^2\sin(\beta y+\gamma)] \\ -a\alpha(b^2+\beta^2)\sin(bx+c)\sin(\beta y+\gamma) = 0. \end{aligned}$$

From this relation we get

$$1°\ a^2b^2 + \alpha^2\beta^2 = 0. \qquad (\text{Put } bx + c = 0,\ \beta y + \gamma = 0).$$

$$2°\ d + \delta = 0. \qquad (\text{Put } bx + c = 0).$$

$$3°\ b^2 + \beta^2 = 0. \qquad (\text{Put } x \text{ and } y \text{ arbitrary}).$$

Thus

$$R(z) = a\,[\sin(bx + c) + \sin(biy + \gamma)]. \tag{26}$$

If we remember that $R(z)$ is positive for $x = 0$, y close to y_0, and for $y = 0$, x close to x_0, we conclude that ab and ac are real, and, moreover, a, b, c either all real or all purely imaginary. If we take b real, we find further $c = (4m + 3)\pi/2$, $\gamma = (4n + 1)\pi/2$. Then (26) takes the form

$$\begin{aligned} R(z) &= +|a|\,[-\cos bx + \cosh by] \\ &= 2|a|\left[\sin^2\left(\frac{b}{2}x\right) + \sinh^2\left(\frac{b}{2}y\right)\right], \end{aligned} \tag{27}$$

or the same result as in formula (20).

Finally, it is possible to gain some information concerning the possible singular points of analytic solutions of (3), supposed single-valued within their domain of existance, without solving the equation. The following principle will help us to discuss the singular points.

A solution of (3) *such that every simply connected region in the plane, no matter how small, contains an interior region, also simply connected, in which the solution is single-valued and analytic, can not have a finite singular point* $z_0 = x_0 + iy_0$, *such that it is possible to find a monotonic point-set* $x_1, x_2 \ldots x_n \ldots$ *or* $y_1, y_2, \ldots, y_n \ldots$ *with* $\lim_{n\to\infty} x_n = x_0$ *or* $\lim_{n\to\infty} y_n = y_0$ *and* $\lim_{n\to\infty} f(x_n) = \infty$ *or* $\lim_{n\to\infty} f(y_n) = \infty$. *In special no limit is possible at the origin other than* 0.

As the function $f(z)$ exists in the whole plane the line $x = x_0$ can not be a singular line throughout. Hence we can find a value $z = x_0 + iY_0$ in the neighbourhood of which $f(z)$ is analytic. But

$$\lim_{n\to\infty} |f(x_n + iY_0)|^2 = \lim_{n\to\infty} |f(x_n)|^2 + |f(iY_0)|^2,$$

which leads to a contradiction, no matter what value $f(iY_0)$ has. The special statement for the origin is proved in the same manner.

This principle excludes poles for the solutions. It also excludes isolated essentially singular points, because we can always find a Jordan curve along which $f(z)$ tends to the limit ∞.* We project the curve on the axes. $f(z)$ can not remain bounded on these projections, hence we can pick out a point set as required. The extension to non-isolated points is possible as long as $f(z)$ can not remain bounded on its singular points. Hence the principle is applicable if the set of singular points is enumerable but generally not if the set is perfect.†

* Proof by Valiron: Démonstration de l'existance pour les fonctions entières, de chemins de détermination infinie, Paris. C. R., 166 (1918), pp. 382–384, for entire functions. The extension to general essentially singular points is obvious. Also W. Gross: Über die Singularitäten analytischer Funktionen, Monatsh. Math. Phys. 29 (1918), pp. 3–47.

† Examples of single-valued and analytic functions whose singular points form a perfect discontinuous set but which remain less than a constant have been constructed by Denjoy and Pompéju.

STOCKHOLM,
August 15, 1922.

ON THE POTENTIAL OF A HOMOGENEOUS SPHERICAL CAP, OF A MAGNETIC SHELL, AND OF A STEADY CIRCULAR CURRENT.

BY C. DE JANS.

1. The potential of a homogeneous spherical cap has been expanded by Maxwell in an infinite series of Legendre's harmonics.* It is my purpose in this little paper to represent that potential by formulae in finite terms, following a method similar to that which Mr. E. Mathy has used in the case of a plane, homogeneous, circular disc.†

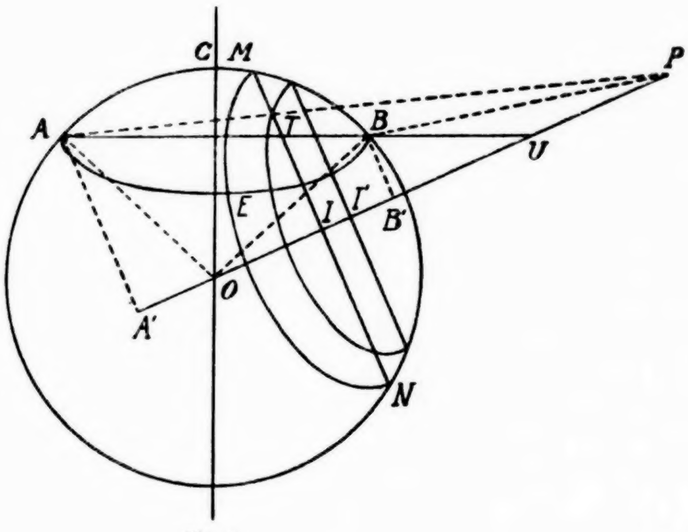

Fig. 1.

Let $A\ E\ B\ C$ (fig. 1) be the given cap, belonging to a sphere with centre O and radius R; the cap is bounded by the small circle $A\ E\ B$, of which C is the pole situated in the cap. We suppose it to be charged with a single layer of an agent which acts according to Newton's law, and whose surface-density is the same in every point. We assume this density to be unity, and desire to

* J. Clerk Maxwell, A Treatise on electricity and magnetism, 2d ed., Oxford 1881, vol. 2, art. 694 (pp. 304–305).

† E. Mathy, Sur le potentiel d'un disque circulaire homogène. Nieuw Archief voor Wiskunde, 2d ser., vol. 10.

calculate the potential of the layer at an arbitrary point P, defined by the distance $O\,P = r$ and the angle $C\,O\,P = \theta$.

If we bear in mind that the potential of a complete sphere, covered with a uniform superficial charge, is constant in the inner space, and has an elementary expression at any external point, it is evident that we may limit our considerations to the case of a cap at the utmost equal to a hemisphere; for in the opposite case the potential will be the difference between the potentials of a complete sphere and of a cap smaller than a hemisphere.

2. Let the cap be determined by the radius R of the sphere and by the angle $C\,O\,A = \alpha$; according to the remark just made, we may suppose

$$0 < \alpha \leqq \tfrac{1}{2}\pi.$$

It follows from symmetry, that the equipotentials are surfaces of revolution with $O\,C$ as their common axis; the point P will therefore be sufficiently determined by the polar coordinates r, θ, introduced above.

We regard r as a magnitude which cannot be negative, and make θ vary within the interval from 0 to π only; if θ were comprised between π and $2\,\pi$, we should have the preceding case by inverting the figure about the axis $O\,C$.

Draw the plane $P\,O\,C$ cutting the sphere in the great circle $A\,C\,B$ and the basis-circle of the cap in the diameter $A\,B$. Let A', B' be the orthogonal projections of A and B on the straight line $O\,P$; U, the intersection of $O\,P$ and $A\,B$; I, I' two points on $O\,P$, taken arbitrarily between A' and B' except as to their mutual distance, which we suppose infinitely small.

Through I and I', draw two planes perpendicular to $O\,P$. These planes intersect the sphere in two small circles, limiting an infinitesimal ring, all the points of which are at the same distance D from P. We take the part of this ring, which belongs to the given cap, as the differential element dS of the latter; its contribution to the potential at P is dS/D.

Let $M\,N$ be the perpendicular to $O\,P$ through the point I in the plane $C\,O\,P$; T, its point of intersection with $A\,B$; M, N, its intersections with the great circle $A\,C\,B$; $M\,E\,N$, the small circle with diameter $M\,N$; E, one of its intersections with the basis-circle of the cap; ψ, the angle $E\,I\,M$; x, the distance $O\,I$, reckoned positive in the sense from O to P; dx, the distance $I\,I'$. Then obviously

$$dS = 2\,\psi\,R\,dx,$$

and, in the triangle $M\,O\,P$,

$$D^2 = R^2 + r^2 - 2\,r\,x. \tag{1}$$

From these relations we derive

$$\frac{dS}{D} = -\frac{2R}{r}\,\psi\, d\, D. \tag{2}$$

3. Let D_1 represent the distance PA, D_2 the distance PB. We have $D_1 > D_2$.

If the indefinite straight line OP does not meet the spherical cap, we will obtain the value of the potential in P, by integrating the formula (2) between the limits D_1 and D_2. This potential is therefore, when $\alpha < \theta < \pi - \alpha$,

$$V = -\frac{2R}{r}\int\limits_{D_1}^{D_2} \psi\, d\, D = \frac{2R}{r}\int\limits_{D_2}^{D_1} \psi\, d\, D. \tag{3}$$

If, on the contrary, the indefinite right line OP does intersect the given cap, the integration of equation (2) between the given limits supplies only a part of the required potential. For clearly this integration leaves out of consideration a small cap, which is a part of the given one; its basis-circle touches the circle AEB, and its vertex is the point Q where the line OP cuts the surface of the given cap. In fig. 2 this small cap is represented by its vertex and its basis-circle FGB, for the case $0 < \theta < \alpha$. In order now to exhaust the whole area of the given cap, we must still take into account the infinitesimal rings corresponding to the distances D, comprised between the limits $PF = D_2$ and $PQ = D_3$; as a consequence, the formula (3) has to be completed by a term V_2, determined by the equation

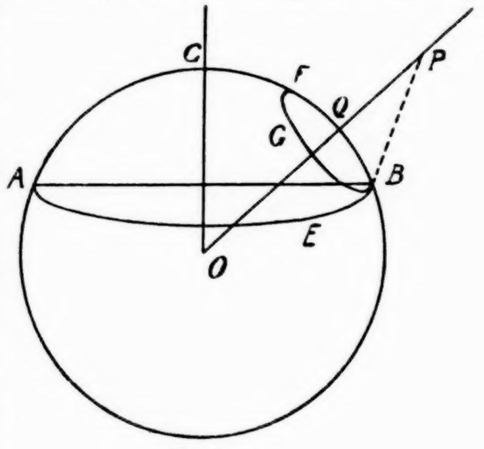

Fig. 2.

$$V_2 = -\frac{2R}{r}\int\limits_{D_4}^{D_3} \psi\, d\, D.$$

so that the whole potential in P becomes

$$V = V_1 + V_2,$$

V_1 denoting the function (3).

Now, if D lies between D_2 and D_3, the corresponding element dS is equal to the total area of its infinitesimal ring; we have therefore to assume $\psi = \pi$, and consequently, when $0 < \theta < \alpha$,

$$(4) \qquad V_2 = \frac{2\pi R}{r}(D_2 - D_3).$$

This formula for the potential of the remaining small spherical cap at a point on its axis, can be easily demonstrated by other modes of reasoning. In the centre of the sphere, i. e. for $r = 0$, it furnishes the value

$$(5) \qquad V = 4\pi R \sin^2 \frac{\alpha}{2}.$$

Replacing V_1 and V_2 by their expressions, we get, with the hypothesis $0 < \theta < \alpha$,

$$(6) \qquad V = \frac{2R}{r}\int_{D_4}^{D_1} \psi \, dD + \frac{2\pi R}{r}(D_2 - D_3).$$

In the same way, we obtain, when $\pi - \alpha < \theta < \pi$,

$$(7) \qquad V_2 = \frac{2\pi R}{r}(D_3 - D_1),$$

and therefore

$$(8) \qquad V = \frac{2R}{r}\int_{D_4}^{D_1} \psi \, dD + \frac{2\pi R}{r}(D_3 - D_1),$$

where D_3 still means the absolute distance between the point P and the point Q, the intersection of the indefinite straight line OP with the given spherical cap.

4. Integration by parts gives

$$\int\limits_{D_2}^{D_1} \psi dD = [\psi D]_{D_2}^{D_1} + \int\limits_{D=D_1}^{D=D_2} D d\psi. \tag{9}$$

In the case $\alpha < \theta < \pi - \alpha$ the angle ψ vanishes at both the integration limits; hence the formula (3) becomes

$$V = \frac{2R}{r} \int\limits_{D=D_1}^{D=D_2} D d\psi. \tag{10}$$

In the case $0 < \theta < \alpha$ the angle ψ vanishes for $D = D_1$, and becomes equal to π for $D = D_2$; so from (6) we have

$$V = \frac{2R}{r} \int\limits_{D=D_1}^{D=D_2} D d\psi - \frac{2\pi R}{r} D_3. \tag{11}$$

In the case $\pi - \alpha < \theta < \pi$ we are led in like manner to the equation

$$V = \frac{2R}{r} \int\limits_{D=D_1}^{D=D_2} D d\psi + \frac{2\pi R}{r} D_3, \tag{12}$$

as a consequence of equation (8).

Before passing to the extreme cases $\theta = \alpha$ and $\theta = \pi - \alpha$, it will be useful to calculate the angle ψ. Let us put $OU = m$, this quantity being reckoned positive in the sense from O to P; then we find, observing that ET and MN are perpendicular to each other,

$$\cos\psi = \frac{(m-x)\cot\theta}{\sqrt{R^2-x^2}}, \tag{13}$$

$$m\cos\theta = R\cos\alpha. \tag{14}$$

so that the whole potential in P becomes

$$V = V_1 + V_2,$$

V_1 denoting the function (3).

Now, if D lies between D_2 and D_3, the corresponding element dS is equal to the total area of its infinitesimal ring; we have therefore to assume $\psi = \pi$, and consequently, when $0 < \theta < \alpha$,

$$V_2 = \frac{2\pi R}{r}(D_2 - D_3). \tag{4}$$

This formula for the potential of the remaining small spherical cap at a point on its axis, can be easily demonstrated by other modes of reasoning. In the centre of the sphere, i. e. for $r = 0$, it furnishes the value

$$V = 4\pi R \sin^2 \frac{\alpha}{2}. \tag{5}$$

Replacing V_1 and V_2 by their expressions, we get, with the hypothesis $0 < \theta < \alpha$,

$$V = \frac{2R}{r}\int_{D_2}^{D_1} \psi \, dD + \frac{2\pi R}{r}(D_2 - D_3). \tag{6}$$

In the same way, we obtain, when $\pi - \alpha < \theta < \pi$,

$$V_2 = \frac{2\pi R}{r}(D_3 - D_1), \tag{7}$$

and therefore

$$V = \frac{2R}{r}\int_{D_2}^{D_1} \psi \, dD + \frac{2\pi R}{r}(D_3 - D_1), \tag{8}$$

where D_3 still means the absolute distance between the point P and the point Q, the intersection of the indefinite straight line OP with the given spherical cap.

4. Integration by parts gives

$$\int\limits_{D_2}^{D_1} \psi \, dD = [\psi D]_{D_2}^{D_1} + \int\limits_{D=D_1}^{D=D_2} D \, d\psi. \tag{9}$$

In the case $\alpha < \theta < \pi - \alpha$ the angle ψ vanishes at both the integration limits; hence the formula (3) becomes

$$V = \frac{2R}{r} \int\limits_{D=D_1}^{D=D_2} D \, d\psi. \tag{10}$$

In the case $0 < \theta < \alpha$ the angle ψ vanishes for $D = D_1$, and becomes equal to π for $D = D_2$; so from (6) we have

$$V = \frac{2R}{r} \int\limits_{D=D_1}^{D=D_2} D \, d\psi - \frac{2\pi R}{r} D_2. \tag{11}$$

In the case $\pi - \alpha < \theta < \pi$ we are led in like manner to the equation

$$V = \frac{2R}{r} \int\limits_{D=D_1}^{D=D_2} D \, d\psi + \frac{2\pi R}{r} D_2, \tag{12}$$

as a consequence of equation (8).

Before passing to the extreme cases $\theta = \alpha$ and $\theta = \pi - \alpha$, it will be useful to calculate the angle ψ. Let us put $OU = m$, this quantity being reckoned positive in the sense from O to P; then we find, observing that ET and MN are perpendicular to each other,

$$\cos \psi = \frac{(m - x) \cot \theta}{\sqrt{R^2 - x^2}}, \tag{13}$$

$$m \cos \theta = R \cos \alpha. \tag{14}$$

The supposition $\theta = \alpha$ gives $m = R$, and therefore

$$\cos\psi = \sqrt{\frac{R-x}{R+x}}\cot\alpha;$$

thus the angle ψ vanishes for $D = D_1$, and becomes equal to $\frac{\pi}{2}$ for $D = D_2$. Hence

$$[\psi D]_{D_2}^{D_1} = -\frac{\pi}{2}D_2,$$

so that, according to (9), the formula (3) becomes

$$(15) \qquad V = \frac{2R}{r}\int_{D=D_1}^{D=D_2} D\,d\psi - \frac{\pi R}{r}D_2.$$

Similarly, when $\theta = \pi - \alpha$,

$$(16) \qquad V = \frac{2R}{r}\int_{D=D_1}^{D=D_2} D\,d\psi + \frac{\pi R}{r}D_1.$$

5. We have now to reduce the preceding expressions of V to known functions. Beginning with (10), we transform it, by means of (1), to

$$\frac{r}{2R}V = \int_{D=D_1}^{D=D_2}\sqrt{R^2+r^2-2rx}\,d\psi.$$

Taking x as the independant variable, the limits of integration will be

$$x' = OA', \qquad x'' = OB',$$

whilst from (13) we deduce

$$(17) \qquad d\psi = \frac{(R^2-mx)\cot\theta\,dx}{(R^2-x^2)\sqrt{R^2-x^2-(m-x)^2\cot^2\theta}}.$$

The substitution of this value in the last formula leads to the equation

$$(18) \qquad \frac{r\tan\theta}{2R}V = \int_{x'}^{x''} \frac{R^2-mx}{R^2-x^2}\sqrt{\frac{R^2+r^2-2rx}{R^2-x^2-(m-x)^2\cot^2\theta}}\,dx,$$

which shews that the potential is expressed by an elliptic integral.

6. In order to perform the integration, we use the notation of Weierstrass. Suppose r to differ from zero, and let u be a new variable, connected with x by the equation

$$(19) \qquad x = \frac{R^2+r^2+4Rr\cos\alpha\cos\theta}{6r} + \wp u,$$

the $\wp$-function being defined by the real constants

$$(20) \qquad \begin{cases} e_1 = \dfrac{R^2+r^2-2Rr\cos\alpha\cos\theta}{3r}, \\ e_3 = -\dfrac{R^2+r^2-2Rr\cos\alpha\cos\theta}{6r} + R\sin\alpha\sin\theta, \\ e_2 = -\dfrac{R^2+r^2-2Rr\cos\alpha\cos\theta}{6r} - R\sin\alpha\sin\theta, \end{cases}$$

whose sum is evanescent and which satisfy the condition $e_1 > e_3 > e_2$.

Now it follows from the consideration of the triangles POA and POB, that

$$(21) \qquad \begin{aligned} 2(R^2+r^2-2Rr\cos\alpha\cos\theta) &= D_1^2+D_2^2, \\ 4Rr\sin\alpha\sin\theta &= D_1^2-D_2^2; \end{aligned}$$

hence the equations (20) at once take the form

$$(22) \qquad e_1 = \frac{D_1^2+D_2^2}{6r}, \qquad e_3 = \frac{D_1^2-2D_2^2}{6r}, \qquad e_2 = \frac{D_2^2-2D_1^2}{6r}.$$

The $\wp$-function here introduced admits a real primitive period $2\omega_1$, and a purely imaginary one $2\omega_2$; we define a third constant ω_3 by the equation

$$\omega_1+\omega_2+\omega_3 = 0.$$

7. To the values $x = x'$, $x = x''$ correspond, by equation (19), two values u', u'' of u; and when we observe that

$$x' = R\cos(\theta + \alpha), \qquad x'' = R\cos(\theta - \alpha), \tag{23}$$

we obtain, in virtue of (20),

$$\wp u' = e_2, \qquad \wp u'' = e_3, \tag{24}$$

so that, leaving multiples of the periods out of account,

$$u' = \omega_2, \qquad u'' = -\omega_3. \tag{25}$$

These relations prove that, within the limits of the integration, u has the form $\omega_2 + t$, where t is real and variable between 0 and ω_1; hence $\wp' u$ is real and positive.

Therefore

$$\frac{dx}{\sqrt{(R^2 + r^2 - 2rx)[R^2 - x^2 - (m - x)^2 \cot^2\theta]}} = \frac{\sqrt{2}\sin\theta}{\sqrt{r}}\, du. \tag{26}$$

Moreover

$$\frac{(R^2 - mx)(R^2 + r^2 - 2rx)}{R^2 - x^2} = -2mr + \frac{A}{x - R} + \frac{B}{x + R}, \tag{27}$$

where

$$A = \frac{(r - R)^2 (m - R)}{2}, \qquad B = \frac{(r + R)^2 (m + R)}{2},$$

or

$$A = \frac{R(r - R)^2}{\cos\theta} \sin\frac{\theta + \alpha}{2} \sin\frac{\theta - \alpha}{2}, \quad B = \frac{R(r + R)^2}{\cos\theta} \cos\frac{\theta + \alpha}{2} \cos\frac{\theta - \alpha}{2}. \tag{28}$$

If we define two constants v, w by the equations

$$\begin{aligned} \wp v &= R - \frac{R^2 + r^2 + 4Rr\cos\alpha\cos\theta}{6r}, \\ \wp w &= -R - \frac{R^2 + r^2 + 4Rr\cos\alpha\cos\theta}{6r}, \end{aligned} \tag{29}$$

we obtain, from (19),

$$(30)\qquad x-R = \wp u-\wp v, \qquad x+R = \wp u-\wp w,$$

so that (27) becomes

$$(31)\qquad \frac{(R^2-mx)(R^2+r^2-2rx)}{R^2-x^2} = -2mr+\frac{A}{\wp u-\wp v}+\frac{B}{\wp u-\wp w}.$$

By this equation, together with (25) and (26), the formula (18) is changed into

$$(32)\qquad \left(\frac{r}{2}\right)^{\frac{3}{2}}\frac{V}{R\cos\theta} = \int_{\omega_3}^{-\omega_3}\left(-2mr+\frac{A}{\wp u-\wp v}+\frac{B}{\wp u-\wp w}\right)du.$$

8. Now from (21) and (29) we deduce

$$6r\wp v = D_1^2+D_2^2-3(r-R)^2, \qquad 6r\wp w = D_1^2+D_2^2-3(r+R)^2,$$

or, according to (22)

$$(33)\qquad \wp v = e_1-\frac{(r-R)^2}{2r}, \qquad \wp w = e_1-\frac{(r+R)^2}{2r}.$$

The last formula shows that $\wp w-e_1$, $\wp w-e_2$, $\wp w-e_3$ do not vanish, so that w is not a semi-period of the $\wp$-function. For this reason we have

$$\int_{\omega_3}^{-\omega_3}\frac{du}{\wp u-\wp w} = -\frac{1}{\wp' w}\left[\log\frac{\sigma(u+w)}{\sigma(u-w)}-2u\zeta w\right]_{\omega_3}^{-\omega_3}.$$

But, by the formulae for the addition of a semi-period,

$$\frac{\sigma(-\omega_3+w)}{\sigma(-\omega_3-w)} = e^{-2\eta_3 w}, \qquad \frac{\sigma(\omega_2+w)}{\sigma(\omega_2-w)} = e^{2\eta_2 w};$$

hence

$$\log\frac{\sigma(-\omega_3+w)}{\sigma(-\omega_3-w)} = -2\eta_3 w+2Mi\pi, \qquad \log\frac{\sigma(\omega_2+w)}{\sigma(\omega_2-w)} = 2\eta_2 w+2Mi\pi,$$

where M denotes any arbitrary integer. So we are led to the equation

$$(34)\qquad \int_{\omega_3}^{-\omega_3} \frac{du}{\wp u - \wp w} = -\frac{2}{\wp' w}(\eta_1 w - \omega_1 \zeta w).$$

9. Excluding for the moment the case $r = R$, the first relation (33) shows that v is not a semi-period of the $\wp$-function; therefore we have also

$$(35)\qquad \int_{\omega_3}^{-\omega_3} \frac{du}{\wp u - \wp v} = -\frac{2}{\wp' v}(\eta_1 v - \omega_1 \zeta v);$$

and now equation (32) is transformed to

$$(36)\qquad \left(\frac{r}{2}\right)^{\frac{3}{2}} \frac{V}{R\cos\theta} = -2mr\omega_1 - \frac{2A}{\wp' v}(\eta_1 v - \omega_1 \zeta v) - \frac{2B}{\wp' w}(\eta_1 w - \omega_1 \zeta w).$$

The value of x being comprised between $-R$ and $+R$, we have

$$x - R < 0, \qquad x + R > 0;$$

hence it is plain, from the relations (30), that

$$\wp v > \wp u > \wp w.$$

Thus the constants $\wp v$ and $\wp w$ are outside the range of variation of the function $\wp u$. But we know, by the equations (24), that $\wp u$ varies between e_3 and e_2; therefore

$$\wp v > e_3, \qquad \wp w < e_2.$$

Since, on the other hand, in virtue of the first equation (33), $\wp v - e_1$ is negative, we conclude

$$e_3 < \wp v < e_1, \qquad \wp w < e_2.$$

These inequalities shew that v is of the form $\omega_1 + it$, t being a real quantity comprised between 0 and ω_2/i, and that w is a purely imaginary magnitude

whose modulus lies between the same limits. If follows at the same time that $\wp' v$ is a positive pure imaginary, $\wp' w$ a negative one.

To calculate $\wp' v$ and $\wp' w$, we apply the formula

$$(\wp' z)^2 = 4(\wp z - e_1)(\wp z - e_2)(\wp z - e_3).$$

The values of $\wp v - e_1$, $\wp w - e_1$ are got at once from (33). Making $u = \omega_2$ in (30), and taking (23) and (25) into account, we obtain further

$$(37)\qquad \begin{aligned} \wp v - e_2 &= \quad R - x' = \quad 2R\sin^2\frac{\theta+\alpha}{2}, \\ \wp w - e_2 &= -R - x' = -2R\cos^2\frac{\theta+\alpha}{2}; \end{aligned}$$

whilst for $u = -\omega_3$ we have

$$(38)\qquad \begin{aligned} \wp v - e_3 &= \quad R - x'' = \quad 2R\sin^2\frac{\theta-\alpha}{2}, \\ \wp w - e_3 &= -R - x'' = -2R\cos^2\frac{\theta-\alpha}{2}. \end{aligned}$$

When these values are substituted in the above formula for $(\wp' z)^2$, it gives

$$(39)\qquad \begin{aligned} \wp' v &= i\frac{2\sqrt{2}}{\sqrt{r}}R|r-R|\sin\frac{\theta+\alpha}{2}\sin\frac{\theta-\alpha}{2}, \\ \wp' w &= -i\frac{2\sqrt{2}}{\sqrt{r}}R(r+R)\cos\frac{\theta+\alpha}{2}\cos\frac{\theta-\alpha}{2}, \end{aligned}$$

and so, on account of (28),

$$(40)\qquad \frac{2A}{\wp' v} = \frac{|r-R|\sqrt{r}}{i\sqrt{2}\cos\theta}, \qquad \frac{2B}{\wp' w} = -\frac{(r+R)\sqrt{r}}{i\sqrt{2}\cos\theta}.$$

Combining these results with equation (30), the function V finally is given by

(41)
$$V = \frac{2R}{r}\left\{|r-R|\frac{\omega_1\zeta v - \eta_1 v}{i} + (r+R)\frac{\eta_1 w - \omega_1\zeta w}{i} - 2\omega_1 R\sqrt{2r}\cos\alpha\right\},$$

and this equation affords the means of calculating the potential, when the angle θ is subjected to the condition $\alpha < \theta < \pi - \alpha$.

10. In deriving the equation (41), we have excluded (art. 9) the hypothesis $r = R$. When, however, this particular value is assigned to r, we ought to write, instead of (27),

$$\frac{(R^2 - mx)(R^2 + r^2 - 2rx)}{R^2 - x^2} = -2mR + \frac{2R^2(m+R)}{R+x},$$

so that the right-hand side of equation (31) has to be replaced by

$$-2mR + \frac{B}{\wp u - \wp w},$$

with $B = 2R^2(m+R)$. Thus, in the present case, the term containing A disappears from equation (32), and the formula then obtained for the potential is the same as that furnished by (41) for the value $r = R$. Therefore equation (35) is seen to hold for any position of the point P corresponding to an angle θ between α and $\pi - \alpha$, and to a non evanescent value of r.

11. Let us next consider the case $0 < \theta < \alpha$.

The integral in equation (11) has still to be calculated as above. First suppose $r \gtrless R$.

Remembering that $\wp' v$ must be a positive imaginary, we must change the sign of i in the first of the formulae (39), by which change in the first of the equations (40) the sign of the right-hand side is altered. Accordingly, the first term in the second member of (41) has its sign changed, so that we find, when we further consider that $D_3 = |r - R|$,

$$(42)\quad V = \frac{2R}{r}\left\{ |r - R| \left(\frac{\eta_1 v - \omega_1 \zeta v}{i} - \pi\right) + \text{the same terms as in (41)}\right\}.$$

This formula too is easily seen to remain correct for the value $r = R$.

In the third place, when θ satisfies the condition $\pi - \alpha < \theta < \pi$, the potential, given by equation (12), is found to be expressed by the formula (41), where the second term is replaced by

$$(r + R)\left(\frac{\omega_1 \zeta w - \eta_1 w}{i} + \pi\right),$$

and this is still true when r becomes equal to R.

12. When $\theta = \alpha$, we have $m = R$, and therefore, instead of (27),

$$\frac{(R^2 - mx)(R^2 + r^2 - 2rx)}{R^2 - x^2} = -2Rr + \frac{R(r+R)^2}{x+R};$$

whence the right-hand side of (31) becomes, putting $B = R(r+R)^2$,

$$-2Rr + \frac{B}{\wp u - \wp w}.$$

Thus the integral in equation (15) acquires the value

$$\int D\, d\psi = (r+R)\frac{\eta_1 w - \omega_1 \zeta w}{i} - 2\omega_1 R\sqrt{2r\cos\alpha};$$

and, bearing in mind that in (15) $D_2 = |r - R|$, we obtain

$$(43)\quad V = \frac{2R}{r}\left\{-\frac{\pi}{2}|r-R| + (r+R)\frac{\eta_1 w - \omega_1 \zeta w}{i} - 2\omega_1 R\sqrt{2r\cos\alpha}\right\},$$

which formula holds good also for $r = R$.

It is to be noticed that the last equation is a particular case of (41) as well as of (42). For if we attribute to θ the value α, v becomes, according to (38), equal to $-\omega_3$. The coefficient of $|r-R|$ in (41) therefore gets the value

$$\frac{1}{i}(-\omega_1 \zeta \omega_3 + \eta_1 \omega_3) = \frac{1}{i}(-\omega_1 \eta_3 + \eta_1 \omega_3) = \frac{1}{i}(\eta_2 \omega_1 - \eta_1 \omega_2),$$

which is known to equal $-\pi/2$. Thus in the present case the formula (41) coincides with (43). In like manner the coefficient of $|r-R|$ in equation (42) is seen also to become $-\pi/2$, so that (42) in its turn is changed into (43). Accordingly the formulae (41), (42) and (43) agree with the necessary continuity of the potential in space.

Similarly, when $\theta = \pi - \alpha$, it can be shown without difficulty that, for all values of r except perhaps the value zero, the equation (16) leads to the following result:

$$V = \frac{2R}{r}\left\{\frac{\pi}{2}(r+R) + |r-R|\frac{\omega_1 \zeta v - \eta_1 v}{i} - 2\omega_1 R\sqrt{2r}\cos\alpha\right\}.$$

and that this formula constitutes a continuous transition between the equation (41) and the one which is arrived at when $\theta > \pi - \alpha$.

13. Finally it is easily proved, that in the cases $\theta = 0$ and $\theta = \pi$, in which the Weierstrassian functions degenerate, the formula (42) and the corresponding one for the case $\pi - \alpha < \theta < \pi$ supply the elementary expressions (4) and (7) for the potential of a spherical cap at a point on its axis. Again, all the above formulae remain valid even when r is made equal to zero (the case hitherto excluded; vide art. 6); they then lead to the formula (5)*. Lastly, they make the potential vanish at infinity.

14. Having now discussed all the alternatives, we see that, of the formulae given above, only the three following are to be retained:

1° when $0 \leqq \theta \leqq \alpha$,

$$\text{(I)} \quad V = \frac{2R}{r}\left\{\left(\frac{\eta_1 v - \omega_1 \zeta v}{i} - \pi\right)(r - R) + \frac{\eta_1 w - \omega_1 \zeta w}{i}(r + R) - 2\,\omega_1 R \sqrt{2r}\cos\alpha\right\};$$

2° when $\alpha \leqq \theta \leqq \pi - \alpha$,

$$\text{(II)} \quad V = \frac{2R}{r}\left\{\frac{\omega_1 \zeta v - \eta_1 v}{i}(r - R) + \frac{\eta_1 w - \omega_1 \zeta w}{i}(r + R) - 2\,\omega_1 R \sqrt{2r}\cos\alpha\right\};$$

3° when $\pi - \alpha \leqq \theta \leqq \pi$,

$$\text{(III)} \quad V = \frac{2R}{r}\left\{\frac{\omega_1 \zeta v - \eta_1 v}{i}(r - R) + \left(\frac{\omega_1 \zeta w - \eta_1 w}{i} + \pi\right)(r + R) - 2\,\omega_1 R \sqrt{2r}\cos\alpha\right\}.$$

It need hardly be said that, although these formulae are general, in the special cases $r = 0$, $\theta = 0$, $\theta = \pi$ the actual calculations are to be performed by means of the simpler equations given in art. 3.

By subtraction of the potentials of two spherical caps, the potential of a spherical zone will be obtained.

* This fact may be more easily verified in the formulae (I'), (II'), (III') of p. 200, which are equivalent to those hitherto established.

15. All the elements which occur in the expressions just written down, can in each particular case, the point P being given, be calculated numerically by the help of the classical methods of the theory of elliptic functions, based on the use of Legendre's tables or of expansions in infinite series. If we wish to use the tables, the formulae may, however, be thrown into a more convenient form.

Calling the modulus k, and the multiplier g, we have, by equations (22),

$$(44) \qquad k^2 = \frac{e_3 - e_2}{e_1 - e_2} = \frac{D_1^2 - D_2^2}{D_1^2}, \qquad g^2 = e_1 - e_2 = \frac{D_1^2}{2r};$$

and, if K and E be the complete elliptic integrals of the first and second kind for the modulus k,

$$(45) \qquad \omega_1 = \frac{K}{g}, \qquad \eta_1 = gE - \frac{e_1 K}{g}.$$

The complementary modulus will as usual, be denoted by k', and the corresponding complete integrals by K' and E'.

Let us now represent by the symbol $\wp_0$ a new $\wp$-function, defined by the constants

$$(46) \qquad (e_1)_0 = -e_2, \qquad (e_3)_0 = -e_3, \qquad (e_2)_0 = -e_1.$$

which give $(e_1)_0 > (e_3)_0 > (e_2)_0$. If g_2, g_3 be the invariants of our original $\wp$-function, defined by the equations (20) or (22), the invariants of the function $\wp_0$ will be $g_2, -g_3$; and, if we associate a subscript index zero with the different functions of Weierstrass and of Jacobi that can be constructed by means of these last invariants, the well-known formulae expressing the homogeneity of these functions give, z being an independant variable.

$$(47) \qquad \begin{cases} \wp(iz) = \wp(iz; g_2, g_3) = -\wp(z; g_2, -g_3) = -\wp_0(z), \\ \wp'(iz) = i\wp_0'(z), \\ \zeta(iz) = -i\zeta_0(z), \\ \mathrm{sn}(iz) = i\dfrac{\mathrm{sn}_0(z)}{\mathrm{cn}_0(z)}. \end{cases}$$

The constants corresponding to the functions with the index 0 we will also affect with the same index; on so doing we find, in addition to (46), the following relations between the constants of the two sets of functions:*

$$(48)\qquad \begin{aligned} &g = g_0, \qquad k = k'_0, \qquad k' = k_0, \\ &\omega_1 = \frac{(\omega_2)_0}{i}, \qquad \omega_2 = i(\omega_1)_0, \qquad K = K'_0, \qquad K' = K_0. \end{aligned}$$

16. The known formula

$$\mathrm{sn}^2(gv) = \frac{g^2}{\wp v - e_2}$$

becomes, in virtue of the first equation (37) and the second (44),

$$(49)\qquad \mathrm{sn}^2(gv) = \frac{D_1^2}{4\,Rr\sin^2\dfrac{\theta+\alpha}{2}}.$$

The right-hand side is ≥ 1, because

$$(50)\qquad D_1^2 = (r-R)^2 + 4\,Rr\sin^2\frac{\theta+\alpha}{2}.$$

We have observed (art. 9) that v is of the form $\omega_1 + it$; since, on the other hand, ω_1 is known by the first equation (45), we shall know the value of v if we determine that of t. Now

$$\mathrm{sn}^2(gv) = \mathrm{sn}^2(g\omega_1 + igt) = \mathrm{sn}^2(K + igt) = \frac{\mathrm{cn}^2(igt)}{\mathrm{dn}^2(igt)} = \frac{1-\mathrm{sn}^2(igt)}{1-k^2\mathrm{sn}^2(igt)}.$$

Substituting in this equation for the left-hand side its value given by (49), and for $\mathrm{sn}^2(igt)$ its value deduced from the last equation (47), we get

$$\mathrm{sn}_0(gt) = \frac{|r-R|}{D_2}.$$

Therefore, if we define a real angle φ (comprised between 0 and $\pi/2$) by the relation

$$(51)\qquad \sin\varphi = \frac{|r-R|}{D_2},$$

* We define g_0 by $g_0 = +\sqrt{(e_1)_0 - (e_2)_0}$.

we have

$$gt = \int_0^\varphi \frac{d\varphi}{\sqrt{1-k_0^2\sin^2\varphi}}.$$

Hence, seeing that $k_0 = k'$, we find

$$t = \frac{1}{g} F(k', \varphi), \tag{52}$$

and the value of t is thus obtained. It follows from this equation that, according to our former remark, t is real and comprised between 0 and K'/g, or, which is the same thing, between 0 and ω_2/i.

17. We have further

$$\zeta v = \zeta(\omega_1 + it) = \eta_1 + \zeta(it) - \frac{1}{2}\frac{\wp'(it)}{e_1 - \wp(it)},$$

or, in virtue of the equations (46) and (47),

$$\zeta v = \eta_1 - i\zeta_0 t - \frac{i}{2}\frac{\wp_0' t}{\wp_0 t - (e_2)_0}.$$

But, on account of a well known formula of the theory of elliptic functions,

$$\zeta_0 t + \frac{1}{2}\frac{\wp_0' t}{\wp_0 t - (e_2)_0} = g_0 E(k_0, \varphi) - (e_1)_0 t,$$

or, by means of the relations (46) and (48),

$$= gE(k', \varphi) + e_2 t;$$

therefore

$$\zeta v = \eta_1 - igE(k', \varphi) - ie_2 t.$$

From this and the value of v we deduce

$$\frac{\omega_1 \zeta v - \eta_1 v}{i} = -(\eta_1 + \omega_1 e_2)t - KE(k', \varphi).$$

Since the equalities (45) give

$$\eta_1 + \omega_1 e_2 = g(E - K), \tag{53}$$

we have, considering also the equation (52),

$$\frac{\omega_1 \zeta v - \eta_1 v}{i} = (K - E)\, F(k', \varphi) - K E(k', \varphi). \tag{54}$$

18. Again,

$$\operatorname{sn}^2(gw) = \frac{g^2}{\wp w - e_2} = -\frac{D_1^2}{4 Rr \cos^2 \dfrac{\theta + \alpha}{2}}.$$

Now, w being a pure imaginary, we put $w = it'$, and so obtain

$$\left[\frac{\operatorname{sn}_0(gt')}{\operatorname{cn}_0(gt')}\right]^2 = \frac{D_1^2}{4 Rr \cos^2 \dfrac{\theta + \alpha}{2}}.$$

Combining this with (50), we get

$$\operatorname{sn}_0(gt') = \frac{D_1}{r + R};$$

and putting

$$\frac{D_1}{r + R} = \sin \psi, \tag{55}$$

t' is seen to be determined by the equation

$$t' = \frac{1}{g} F(k', \psi), \tag{56}$$

where the angle ψ is real.

This result shows that w is comprised between 0 and iK'/g, i. e. between 0 and ω_2, in accordance with the assertion in art. 9.

Further

$$\zeta w = \zeta(it') = -i\zeta_0 t'. \tag{57}$$

But

$$g_0 E(k_0, \psi) = (e_1)_0 t' + \zeta_0 t' + \frac{1}{2} \frac{\wp_0' t'}{\wp_0 t' - (e_2)_0}.$$

or

(58) $$\zeta_0 t' = g E(k', \psi) + e_2 t' - \frac{1}{2} \frac{\wp'_0 t'}{\wp_0 t' - (e_2)_0}.$$

On the other hand

$$\wp_0 t' - (e_2)_0 = \left[\frac{g}{\mathrm{sn}_0 (g t')}\right]^2,$$

whence, by differentiation,

$$\wp'_0 t' = - \frac{2 g^3 \,\mathrm{cn}_0 (g t') \,\mathrm{dn}_0 (g t')}{[\mathrm{sn}_0 (g t')]^3}.$$

Therefore

$$- \frac{1}{2} \frac{\wp'_0 t'}{\wp_0 t' - (e_2)_0} = g \frac{\mathrm{cn}_0 (g t') \,\mathrm{dn}_0 (g t')}{\mathrm{sn}_0 (g t')};$$

and, since

$$\mathrm{sn}_0 (g t') = \sin \psi, \quad \mathrm{cn}_0 (g t') = \cos \psi, \quad \mathrm{dn}_0 (g t') = \sqrt{1 - k'^2 \sin^2 \psi},$$

we have

$$- \frac{1}{2} \frac{\wp'_0 t'}{\wp_0 t' - (e_2)_0} = g \cot \psi \sqrt{1 - k'^2 \sin^2 \psi},$$

the positive determination being taken for the square root.

Substituting in equation (58) and then in (57), we obtain

$$- \frac{\omega_1 \zeta w}{i} = K E(k', \psi) + \omega_1 e_2 t' + K \cot \psi \sqrt{1 - k'^2 \sin^2 \psi},$$

$g \omega_1$ having been replaced by its value K.

Since $\eta_1 w = i \eta_1 t'$, we further get

$$\frac{\eta_1 w - \omega_1 \zeta \omega}{i} = K E(k', \psi) + (\eta_1 + \omega_1 e_2) t' + K \cot \psi \sqrt{1 - k'^2 \sin^2 \psi},$$

or, by (53) and (56),

$$\frac{\eta_1 w - \omega_1 \zeta w}{i} = K E(k', \psi) + (E - K) F(k', \psi) + K \cot \psi \sqrt{1 - k'^2 \sin^2 \psi}.$$

If in the last term of this formula we substitute for ψ the value deduced from (55), we obtain

(59)

$$\frac{\eta_1 w - \omega_1 \zeta w}{i} = K E(k', \psi) - (K - E) F(k', \psi) + \frac{2\lambda K R r}{(r+R) D_1} (\cos\theta + \cos\alpha),$$

λ being $+1$ when $\theta < \pi - \alpha$, and -1 when $\theta > \pi - \alpha$.

19. Writing, for the sake of brevity,

$$S(z) = K E(k', z) - (K - E) F(k', z), \tag{60}$$

the formulae (54) and (59) become

(61)

$$\frac{\omega_1 \zeta v - \eta_1 v}{i} = -S(\varphi),$$

$$\frac{\eta_1 w - \omega_1 \zeta w}{i} = S(\psi) + \frac{2\lambda K R r}{D_1 (r+R)} (\cos\theta + \cos\alpha),$$

whilst

$$2\omega_1 R \sqrt{2r} \cos\alpha = \frac{4 K R r}{D_1} \cos\alpha. \tag{62}$$

By substitution in the formulae (I), (II), (III) of art. 14, we obtain the following expressions for the potential:

1° When $0 \leqq \theta \leqq \alpha$,

$$V = \frac{2R}{r} \left\{ (r - R)(S\varphi - \pi) + (r+R) S\psi + \frac{2KRr}{D_1} (\cos\theta - \cos\alpha) \right\}; \tag{I'}$$

2° when $\alpha \leqq \theta \leqq \pi - \alpha$,

$$V = \frac{2R}{r} \left\{ -(r - R) S\varphi + (r+R) S\psi + \frac{2KRr}{D_1} (\cos\theta - \cos\alpha) \right\}; \tag{II'}$$

3° when $\pi - \alpha \leqq \theta \leqq \pi$,

$$V = \frac{2R}{r} \left\{ -(r - R) S\varphi - (r+R)(S\psi - \pi) + \frac{2KRr}{D_1} (\cos\theta - \cos\alpha) \right\}. \tag{III'}$$

These formulae agree with one another for the limiting values of θ, as will easily be seen if the relation

$$S\left(\frac{\pi}{2}\right) = KE' + EK' - KK' = \frac{\pi}{2}$$

be kept in mind. They also subsist when r is made equal to zero; in this case they lead to equation (5). When $\theta = 0$ they furnish the relation (4), and when $\theta = \pi$ the relation (7), where of course $D_1 = D_2$. In the general case, after φ and ψ have been determined by the equations (51) and (55), the formulae allow the direct calculation of the potential by means of the tables of elliptic integrals.

20. *Remark.* We may incidentally notice that the form of the expressions (I), (II), (III) of art. 14 undergoes but a very slight change when, instead of the variable u and the $\wp$-function introduced in art. 6, we adopt a variable u_1, bearing to u a real and constant ratio μ, say

$$u_1 = \mu u,$$

and the $\wp$-function with periods $2\mu\omega_1$ and $2\mu\omega_2$. If we characterize by a subscript 1 this new $\wp$-function and the functions of Weierstrass associated therewith, the formulae of homogeneity of these functions give

$$\zeta_1 u_1 = \zeta(\mu u;\, \mu\omega_1,\, \mu\omega_2) = \frac{1}{\mu}\zeta(u;\, \omega_1,\, \omega_2) = \frac{1}{\mu}\zeta u,$$

$$\wp_1 u_1 = \frac{1}{\mu^2}\wp u,$$

etc.; and, when $(e_1)_1$ denotes the greatest and $(e_2)_1$ the least of the e-constants of the function $\wp_1$, and $2(\omega_1)_1$, its real period, we have, since μ^2 is supposed to be positive,

$$(\omega_1)_1 = \mu\omega_1, \qquad (\omega_2)_1 = \mu\omega_2, \qquad (e_h)_1 = \frac{1}{\mu^2}e_h, \qquad (\eta_h)_1 = \frac{1}{\mu}\eta_h, \text{ etc.,}$$

h denoting any one of the numbers 1, 2, 3.

In this way, putting

$$v_1 = \mu v, \qquad w_1 = \mu w,$$

we obtain, instead of (I),

$$V = \frac{2R}{r}\left\{\left[\frac{(\eta_1)_1 v_1 - (\omega_1)_1 \zeta_1 v_1}{i} - \pi\right](r - R)\right.$$

$$\left. + \frac{(\eta_1)_1 w_1 - (\omega_1)_1 \zeta_1 w_1}{i}(r + R) - \frac{2(\omega_1)_1 R\sqrt{2r}}{\mu}\cos\alpha\right\},$$

and similar formulae instead of (II) and (III). By the substitutions considered, the equations (I′), (II′) and (III′) of art. 19 suffer no modification whatever.

21. In order to deduce from the preceding formulae the potential of a uniform magnetic shell in the shape of a spherical cap, or the magnetic potential of a steady circular electric current, we have to perform some preliminary calculations, leading to the following results:

$$\frac{\partial k}{\partial r} = -\frac{k(r^2 - R^2)}{2rD_1^2}, \qquad \frac{\partial k'}{\partial r} = \frac{k^2(r^2 - R^2)}{2k'rD_1^2},$$

$$\frac{\partial K}{\partial r} = -\frac{r^2 - R^2}{2rD_1^2}\cdot\frac{E - k'^2K}{k'^2}, \qquad \frac{\partial E}{\partial r} = \frac{r^2 - R^2}{2rD_1^2}(K - E),$$

$$\frac{\partial(K - E)}{\partial r} = -\frac{r^2 - R^2}{2rD_1^2}\cdot\frac{k^2E}{k'^2},$$

$$\frac{\partial D_1}{\partial r} = \frac{r - R\cos(\theta + \alpha)}{D_1}, \qquad \frac{\partial D_2}{\partial r} = \frac{r - R\cos(\theta - \alpha)}{D_2},$$

and putting

$$\Delta(z) = \sqrt{1 - k'^2\sin^2 z},$$

where the positive determination of the square root is adopted,

$$\frac{\partial S(z)}{\partial r} = \frac{\cos z}{\Delta(z)}\left[\frac{r^2 - R^2}{2rD_1^2}(K - E)\sin z + \frac{E - k'^2K\sin^2 z}{\cos^2 z}\,\frac{\partial \sin z}{\partial r}\right].$$

Putting

$$\varepsilon = +1 \text{ when } r > R, \qquad \varepsilon = -1 \text{ when } r < R,$$

$$\varkappa = +1 \text{ when } \theta > \alpha, \qquad \varkappa = -1 \text{ when } \theta < \alpha,$$

$$\lambda = +1 \text{ when } \theta < \pi - \alpha, \qquad \lambda = -1 \text{ when } \theta > \pi - \alpha,$$

we further have

$$\sin\varphi = \varepsilon\frac{r-R}{D_2},\quad \cos\varphi = \varkappa\frac{2\sqrt{Rr}}{D_2}\sin\frac{\theta-\alpha}{2},\quad \Delta\varphi = \frac{2\sqrt{Rr}}{D_1}\sin\frac{\theta+\alpha}{2},$$

$$\sin\psi = \frac{D_1}{r+R},\quad \cos\psi = \lambda\frac{2\sqrt{Rr}}{r+R}\cos\frac{\theta+\alpha}{2},\quad \Delta\psi = \frac{2\sqrt{Rr}}{r+R}\cos\frac{\theta-\alpha}{2},$$

$$\varepsilon(r-R)\frac{\partial S\varphi}{\partial r} = \varkappa\frac{R(r^2-R^2)E}{D_1 D_2^2}(\cos\alpha-\cos\theta),$$

$$(r+R)\frac{\partial S\psi}{\partial r} = \lambda\frac{R(r-R)K}{D_1(r+R)}(\cos\alpha+\cos\theta),$$

$$\frac{\partial}{\partial r}\left(\frac{Kr}{D_1}\right) = \frac{K}{2D_1} - \frac{r^2-R^2}{2D_1D_2^2}E.$$

We are now in a position to determine the potential Ω of a uniform magnetic shell, which covers the spherical cap hitherto considered. We suppose this shell to have unit strength, its positive side being turned outwards with respect to the sphere.

With these conditions we have*

$$\Omega = -\frac{1}{R}\frac{\partial(rV)}{\partial r}.$$

Applying this formula to the function V defined by the equations (I'), (II') and (III'), and using the foregoing relations, we obtain

1° when $0 \leqq \theta \leqq \alpha$,

$$(\mathrm{I}'')\qquad \Omega = 2\varepsilon(\pi - S\varphi) - 2S\psi + \frac{4KR}{D_1(r+R)}(R\cos\alpha - r\cos\theta);$$

2° when $\alpha \leqq \theta \leqq \pi-\alpha$;

$$(\mathrm{II}'')\qquad \Omega = 2\varepsilon S\varphi - 2S\psi + \frac{4KR}{D_1(r+R)}(R\cos\alpha - r\cos\theta);$$

* J. C. Maxwell, op. cit., vol. 2, art. 670 and 694.

3° when $\pi - \alpha \leqq \theta \leqq \pi$,

$$(\mathrm{III}'')\qquad \Omega = 2\varepsilon S\varphi + 2(S\psi - \pi) + \frac{4KR}{D_1(r+R)}(R\cos\alpha - r\cos\theta).$$

Owing to the relation $S(0) = 0$, the first of these equations implies a discontinuity of the potential when the potentiated point P traverses the shell, i. e. for $r = R$; this discontinuity is measured by 4π, as it ought to be according to a well known property of magnetic shells.

22. The expressions (I''), (II''), (III'') are at the same time those of the magnetic potential of a steady linear current of unit strength, flowing along the circular boundary of the spherical cap. But, as the potential of the current for a given radius of the latter does not depend on the aperture α of the cap, the formulae expressing that potential must be equivalent to one another. We therefore shall use only the last of the three written above and, with a view to simplification, we will choose α equal to $\frac{\pi}{2}$, so that the origin of the radii vectores lies in the centre of the current circuit.*

Then R is the radius of the circular circuit, while D_1 and D_2 still denote the maximum and minimum distance of the potentiated point to that circuit. If, moreover, we suppose the positive side of the shell to coat the inner side of the hemisphere, the expression of the potential, as given by (III''), becomes

$$\Omega = 2\pi - 2\varepsilon S\varphi - 2S\psi + \frac{4KRr\cos\theta}{D_1(r+R)}.$$

Now we may, without altering the value of the potential, replace the hemispherical shell by a flat one coinciding with the plane of the circuit. If, in addition to this, we measure the angle θ from the normal to the positive side of the shell (on which side P is supposed to be situated), we must change the sign of $\cos\theta$, and the formula then becomes

$$(63)\qquad \Omega = 2\pi - 2\varepsilon S(\varphi) - 2S(\psi) - \frac{4KRr\cos\theta}{D_1(r+R)}.$$

In this formula φ and ψ are defined by the equations (51) and (55), the function S by equation (60), k by the first of the equations (44).

* As far as the author is aware, the formulae expressing the potential of a circular current in terms of elliptic integrals, hitherto published, all suppose the origin of the radii vectores to coincide with the potentiated point.

23. In the particular case where the potentiated point belongs to the axis of the circuit, we have $\theta = 0$, and consequently

$$D_1 = D_2 = \sqrt{R^2+r^2}, \quad k = 0, \quad k' = 1, \quad K = E = \frac{\pi}{2}, \quad S(z) = \frac{\pi}{2}\sin z.$$

Then from (63) we have

$$\Omega = 2\pi\left(1 - \frac{\varepsilon}{2}\sin\varphi - \frac{1}{2}\sin\psi - \frac{Rr}{(r+R)\sqrt{r^2+R^2}}\right).$$

But

$$\varepsilon\sin\varphi + \sin\psi = \frac{r-R}{\sqrt{r^2+R^2}} + \frac{\sqrt{r^2+R^2}}{r+R} = \frac{2r^2}{(r+R)\sqrt{r^2+R^2}},$$

whence, calling ξ the parallax of the potentiated point with respect to the circuit,

$$\Omega = 2\pi\left(1 - \frac{r}{\sqrt{r^2+R^2}}\right) = 2\pi(1-\cos\xi).$$

This is the well known elementary formula.

24. Finally, it may be interesting to remark that the equations (I″), (II″) and (III″) can be arrived at directly in the same way that has led us to (I′), (II′), (III′). For if β be the angle formed by the straight lines OM and MP (fig. 1), the expression of the potential of the magnetic shell in P is, according to a known formula*

$$\Omega = \int_{D=D_1}^{D=D_2} \frac{\cos\beta}{D^2}\, dS.$$

Now

$$\cos\beta = \frac{r^2 - R^2 - D^2}{2RD},$$

so that we get, on account of the formula (2),

$$\Omega = \frac{1}{r}\int_{D_1}^{D_2} \frac{D^2 - r^2 + R^2}{D^2}\,\psi\, dD.$$

* Vide, for instance, J. H. Jeans, The mathematical theory of electricity and magnetism. Cambridge 1908, art. 418, p. 366. dS has the same meaning as in Art. 2.

Integrating by parts, this becomes

$$(64)\qquad \Omega = \frac{1}{r}\left[\psi\,\frac{D^2+r^2-R^2}{D}\right]_{D_1}^{D_2} - \frac{1}{r}\int\limits_{D=D_1}^{D=D_2} \frac{D^2+r^2-R^2}{D}\,d\psi.$$

When $\alpha < \theta < \pi - \alpha$, the integration between the limits D_1 and D_2 supplies the total value of the potential.

When $0 < \theta < \alpha$, we must, in order to obtain the whole potential, add to the right-hand member of this equation the term

$$\int\limits_{D=D_2}^{D=D_3} \frac{\cos\beta}{D^2}\,dS = \frac{\pi}{r}\int\limits_{D_2}^{D_3} \frac{D^2-r^2+R^2}{D^2}\,dD - \frac{\pi}{r}\left[\frac{D^2+r^2-R^2}{D}\right]_{D_2}^{D_3};$$

and, when $\pi - \alpha < \theta < \pi$, the term

$$\int\limits_{D=D_3}^{D=D_1} \frac{\cos\beta}{D^2}\,dS = \frac{\pi}{r}\left[\frac{D^2+r^2-R^2}{D}\right]_{D_3}^{D_1}.$$

In the case $\alpha < \theta < \pi - \alpha$, the formula (64) gives

$$\Omega = -\frac{1}{r}\int\limits_{D=D_1}^{D=D_2} \frac{D^2+r^2-R^2}{D}\,d\psi.$$

In the case $0 < \theta < \alpha$ we must add to the right-hand side of this equation the term

$$\Omega_1 = \frac{\pi}{r}\cdot\frac{r^2-R^2+D_3^2}{D_3} = 2\pi\epsilon,$$

where ϵ has the same meaning as in art. 21.

In the case $\pi - \alpha < \theta < \pi$ we have similarly to majorate the right-hand side by the term

$$\Omega_2 = -2\pi.$$

Now, putting

$$G(x) = \sqrt{(R^2 + r^2 - 2rx)\,[R^2 - x^2 - (m - x)^2 \cot^2\theta]},$$

we have, in virtue of our previous equations (1) and (17),

$$(65)\quad \int_{D=D_1}^{D=D_2} \frac{D^2 + r^2 - R^2}{D}\,d\psi = 2r\cot\theta \int_{x'}^{x''} \frac{(R^2 - mx)(r - x)}{R^2 - x^2} \cdot \frac{dx}{G(x)}.$$

But

$$\frac{(R^2 - mx)(r - x)}{R^2 - x^2} = -m + \frac{A'}{x - R} + \frac{B'}{x + R},$$

where

$$A' = \frac{A}{r - R}, \qquad B' = \frac{B}{r + R},$$

A and B being the constants defined in art. 7.

Accordingly, in virtue of (26), (30), (34), and (35), the integral in the right-hand side of equation (65) becomes

$$-\sqrt{\frac{2}{r}}\sin\theta\left[m\,\omega_1 + \frac{2A'}{\wp' v}(\eta_1 v - \omega_1 \zeta v) + \frac{2B'}{\wp' w}(\eta_1 w - \omega_1 \zeta w)\right],$$

so that, according to (40) and remembering equation (14), the left-hand side of (65) becomes, when $\alpha < \theta < \pi - \alpha$,

$$= -2r\left(\frac{\omega_1 \sqrt{2R}\cos\alpha}{\sqrt{r}} + \epsilon\,\frac{\eta_1 v - \omega_1 \zeta v}{i} - \frac{\eta_1 w - \omega_1 \zeta w}{i}\right),$$

and has similar values with the hypotheses $\theta < \alpha$ and $\theta > \pi - \alpha$.

Therefore we have, for $0 \leqq \theta \leqq \alpha$,

$$(\mathrm{I}''')\quad \Omega = -2\epsilon\left(\frac{\eta_1 v - \omega_1 \zeta v}{i} - \pi\right) - 2\,\frac{\eta_1 w - \omega_1 \zeta w}{i} + \frac{2\sqrt{2}\,\omega_1 R\cos\alpha}{\sqrt{r}};$$

again, for $\alpha \leqq \theta \leqq \pi - \alpha$,

$$(\mathrm{II}''')\quad \Omega = -2\epsilon\,\frac{\omega_1 \zeta v - \eta_1 v}{i} - 2\,\frac{\eta_1 w - \omega_1 \zeta w}{i} + \frac{2\sqrt{2}\,\omega_1 R\cos\alpha}{\sqrt{r}};$$

and finally, for $\pi - \alpha \leq \theta \leq \pi$,

$$(\mathrm{III}''') \quad \Omega = -2\varepsilon \frac{\omega_1 \zeta v - \eta_1 v}{i} - 2\left(\frac{\omega_1 \zeta w - \eta_1 w}{i} + \pi\right) + \frac{2\sqrt{2}\,\omega_1 R \cos\alpha}{\sqrt{r}}.$$

These are the expressions of the potential of the magnetic shell in terms of functions of Weierstrass.

Transforming these formulae by means of the equations (61) and (62), we obtain the expressions (I″), (II″), (III″) established above.

By the introduction of the functions and variables of art. 20 in the formulae (I‴), (II‴), and (III‴), the different terms keep their form unchanged, with the exception of the last term of each equation, which is divided by μ as in the article referred to.

ON CYCLIC-HARMONIC CURVES.

By Harold Hilton.

1. In these Annals* R. E. Moritz has given some properties of the curves whose polar equation is

(i) $$r = a\cos p\theta/q + k, \ \ldots,$$

where p and q are positive integers with no common factor, and a, k are positive constants.

A more complete description of these curves is here given, and his results are summarised so far as is necessary to make the extension of his work intelligible.

The curve (1) is unicursal; for $x = r\cos\theta$ and $y = r\sin\theta$ can be expressed rationally in terms of $\tan\theta/2q$.

Equation (1) gives

(ii) $$2\cos p\theta = R^q - c_1 R^{q-2} + c_2 R^{q-4} - \ldots.$$

where $R = 2(r-k)/a$, $c_t = {}^{q-t}C_t + {}^{q-t-1}C_{t-1}$.

This may be put in the form

(iii) $$2\cos p\theta = \alpha r^q + \beta r^{q-1} + \gamma r^{q-2} + \delta r^{q-3} + \ldots + \varkappa r + \lambda,$$

where α, β, γ, δ, ... are polynomials in k whose values are easily calculated. Only odd powers of k are involved in β, δ, ..., and only even powers in α, γ, ...

The finite inflexions are given by

$$(q^2-p^2)a^2\cos^2 p\theta/q + (2q^2+p^2)ak\cos p\theta/q + q^2k^2 + 2p^2a^2 = 0.$$

No inflexions are real, unless $(p^2+q^2)a > q^2k > q^2a$.

Let (x, y) be a real finite focus, and let $c \equiv x+iy$. The equation for c is the condition that the result of eliminating θ between (1) and $re^{\theta i} = c$ shall have equal roots, considered as an equation in r. We find

---

* Vol. 23 (1922), p. 29.

$$(p+q)^p(p-q)^{p-q}\{qk\pm[q^2k^2+(p^2-q^2)a^2]^{\frac{1}{2}}\}^q c^p$$
(iv)
$$= p^p\{pk\pm[q^2k^2+(p^2-q^2)a^2]^{\frac{1}{2}}\}^p a^q, \ldots .$$

where in the ambiguity the two plus or the two minus signs are to be taken.

2. Suppose $k = 0$ and p, q are both odd.

The curve consists of p loops, is of degree $p+q$, and has the symmetry of a regular p-sided polygon.* The origin is a p-ple point,† and there are $\frac{1}{2}p(q-1)$ finite crunodes. The only other point-singularities are at the circular points at infinity.

We shall call the origin in $y^u = x^{u+v}$ a "singularity of type $[u, v]$". It is equivalent to $\frac{1}{2}(u-1)(u+v-3)$ nodes, $u-1$ cusps, $\frac{1}{2}(v-1)(u+v-3)$ bitangents, and $v-1$ inflexions (H. P. A. C., p. 119, Ex. 1).‡

If $p > q$, each circular point is a singularity of type $[q, \frac{1}{2}(p-q)]$, the tangent at the singularity being the line at infinity. The class of the curve is $2p$.

If $q > p$, each circular point is a singularity of type $[\frac{1}{2}(q+p), \frac{1}{2}(q-p)]$, the tangent at the singularity passing through the origin. The class of the curve is $p+q$.

In either case there are p real finite ordinary foci. The equation § 1 (iv) giving the foci reduces to

$$(p+q)^{\frac{1}{2}(p+q)}(p-q)^{\frac{1}{2}(p-q)}c^p = p^p a^p.$$

3. Now suppose $k = 0$, and one of p, q even and the other odd.

The curve consists of $2p$ loops, is of degree $2(p+q)$, and has the symmetry of a regular $2p$-sided polygon. The origin is a $2p$-ple point through which pass $2p$ real linear branches touching each other in pairs. There are $2p(q-1)$ finite crunodes. The only other point-singularities are at the circular points.

If $p > q$, each circular point is a singularity of type $[2q, p-q]$, the tangent at the singularity being the line at infinity. The class of the curve is $4p$.

If $q > p$, each circular point is a singularity of type $[q+p, q-p]$, the tangent at the singularity passing through the origin. The class of the curve is $2(p+q)$.

* If $p = 1$, the curve has one axis of symmetry. If $p = 2$, the curve has the symmetry of the rectangle.

† A multiple point of order p.

‡ References will be made to Hilton's "Plane Algebraic Curves", Clarendon Press, 1920, by means of the letters H. P. A. C.

In either case there are $2p$ real finite ordinary foci. The equation § 1 (iv) reduces to

$$(p+q)^{p+q}(p-q)^{p-q}c^{2p} = p^{2p}a^{2p}.$$

4. Take now the case $k>0$. The properties of the curve may be deduced from those of the curve with $k=0$ by the methods of H. P. A. C., Ch. XI, § 9.

For all values of p and q the curve is of degree $2(p+q)$ and has the symmetry of a regular p-sided polygon; while the origin is a $2p$-ple point.

First suppose p and q both odd.

The rationalised form of the Cartesian equation is got by writing § 1 (iii) in the shape

$$\text{(i)} \quad (-2r^p \cos p\theta + \alpha r^{p+q} + \gamma r^{p+q-2} + \dots + \varkappa r^{p+1})^2$$
$$= (\beta r^{p+q-1} + \delta r^{p+q-3} + \dots + \lambda r^p)^2.$$

The $2p$ linear branches through the origin have distinct tangents.*

As is usually the case for a conchoid of an algebraic curve, equation (i) is not altered by changing k into $-k$.

To obtain the nature of the curve at a circular point, we replace $re^{\theta i}$ and $re^{-\theta i}$ in (i) by x/y and $1/y$. We get thus a projection of the curve in which a circular point has been projected into the origin and the line at infinity into the axis of x, (H. P. A. C., Ch. I, § 3). We can now apply the methods of H. P. A. C., Ch. VI, § 1.

If $p>q$, we find that there is at each circular point a single superlinear branch of order $2q$ with the line at infinity as tangent. This superlinear branch is highly specialised. Its nature is that of the singularity at the origin given by the expansion

$$y = \alpha^{1/q}x^{(p+q)/2q} \pm A\alpha^{2/q}x^{(2p+q)/2q} + B\alpha^{3/q}x^{(3p+q)/2q} \pm C\alpha^{4/q}x^{(4p+q)/2q} + \dots,$$

where each of the q values of $\alpha^{1/q}$ is taken, and all the plus or all the minus signs.

Similarly if $q>p$.

5. Now suppose p is even and q odd.

The rationalised form of the Cartesian equation is got by writing § 1 (iii) in the shape

* We suppose in §§ 4, 5 that the ratio k/a is general. The results must be modified if, for instance, $k=a$.

$$\text{(i)}\quad \begin{aligned}&(-2r^p\cos p\theta+\beta r^{p+q-1}+\delta r^{p+q-3}+\cdots+\lambda r^p)^2\\ &\qquad=(\alpha r^{p+q}+\gamma r^{p+q-2}+\cdots+\varkappa r^{p+1})^2.\end{aligned}$$

The $2p$ linear branches through the origin touch in pairs.

Equation (i) is altered by changing the sign of k. This is exceptional for a conchoid. It is paralleled by the case of a conchoid of a circle with respect to its centre.

The nature of the curve at a circular point is the same as for the curve of § 3.

The above statements hold good for p odd and q even, except that (i) is replaced by

$$\begin{aligned}&(-2r^p\cos p\theta+\beta r^{p+q-1}+\delta r^{p+q-3}+\cdots+\varkappa r^{p+1})^2\\ &\qquad=(\alpha r^{p+q}+\gamma r^{p+q-2}+\cdots+\lambda r^p)^2.\end{aligned}$$

and the $2p$ linear branches through the origin have distinct tangents.

MULTIPLE INTEGRALS IN n-SPACE.

BY PHILIP FRANKLIN.

The purpose of this paper is to give an exposition of the fundamental theorems on integrals in hyper-space. While these theorems go back to Poincaré,* the presentation here given is somewhat simpler than his, and may be of interest in view of the applications of these theorems to modern theoretical physics.

1. **Ordinary integrals in Euclidean space.** We shall take as our starting point the notion of an ordinary multiple integral (of the nth order) in Euclidean n-space. That is, given an n-region defined by a single relation of the form:

$$f(x_1 \cdots x_n) < 0; \tag{1}$$

involving the n special co-ordinates $x_1 \cdots x_n$, and a function of position

$$A(x_1 \cdots x_n) \tag{2}$$

defined inside this n-region, we set up the expression:

$$\int n \int A(x_1 \cdots x_n)\, dx_1 \cdots dx_n \tag{3}$$

which represents the limit (shown to exist under suitable conditions on A) of a certain n-tuple summation precisely analogous to that defining double integrals in the plane or triple integrals in 3-space.

In defining such integrals, no sign is assigned to the "volume" element, which is taken as always positive, so that, in any n-region in which A preserves sign, the integral also has this sign. This convention leads to no difficulties as long as all our integrals are of the same order as the space we are working in, and we use a single fixed co-ordinate system in this space. However, if

* Acta, vol. 9 p. 321.

we change our co-ordinates, we are led to write as the equivalent of (3), referred to the co-ordinate system $y_1 \cdots y_n$:*

$$(4) \qquad \int^n \int \bar{A}(y_1 \cdots y_n) \left| \frac{D(x_1 \cdots x_n)}{D(y_1 \cdots y_n)} \right| dy_1 \cdots dy_n,$$

taken over the n-region given by

$$(5) \qquad \bar{f}(y_1 \cdots y_n) < 0,$$

where the barred functions denote the values of the functions in terms of the new variables and the additional factor in (4) is the absolute value of the Jacobian of the old variables with respect to the new ones. Since the sign of this Jacobian depends on the relation of the senses of the two co-ordinate systems, the absolute value has the effect of changing the sign or not according as the new co-ordinate system is of the opposite or of the same sense as the old.

2. **Oriented cells and regions.** Before proceeding to a generalization of the above definition of an integral, we shall have to discuss the regions of integration with which we shall be concerned. In a Euclidean n-space, with co-ordinates $u_1 \cdots u_n$, an expression of the form:

$$(6) \qquad f(u_1 \cdots u_n) < 0$$

in general defines a connected portion of the space; if there exists a continuous one-to-one point transformation of this portion of the space into the interior of the sphere.

$$(7) \qquad \bar{u}_1^2 + \ldots + \bar{u}_n^2 - 1 < 0$$

our region is an n-cell.

The n-cell may be *oriented* or sensed by assigning a definite order to the co-ordinates in terms of which it is defined. Two n-cells defined by the same function, but with different orders of the variables, are of the same or opposite sense according as the permutation of the variables required to go from one to the other is even or odd.

---

* Scott and Mathews, Determinants, p. 172.

In a space of $m (m > n)$ dimensions, an oriented n-cell is given by a set of equations

$$
(8) \qquad \begin{gathered} x_i = x_i(u_1 \ldots u_n) \qquad 0 < i < m \\ f(u_1 \ldots u_n) < 0 \end{gathered}
$$

together with an order of the u's, the equations (8) serving to map this cell on one in the u-space.

By our definition, all our n-cells may be mapped on a sphere in n-space (7). If we assume the necessary partial derivatives exist and are continuous (as we shall for all cells here used) and construct the Jacobian

$$
(9) \qquad \frac{D(u_1 \ldots u_n)}{D(\bar{u}_1 \ldots \bar{u}_n)},
$$

the order of the u's being that given, and that of the $\bar{u}$'s one fixed on once for all, we may associate a definite sign, that of (9), with each sense of our cell. The sign of (9) is the same for all points of the cell, since it is never zero (as the transformation is one to one) and is continuous.

Our region of integration will be one given by a set of equations of the form (8), except that the function f therein need not correspond to a single cell. We shall, however, assume that, for each point of our region, there may be found a cell which contains the point and all of whose points belong to the region; it will be given by a different inequality in the u's. We further assume that the sense of all these cells as given by the order of the u's and (9) is the same; this restricts us to "orientable regions".* Finally, our regions will be so chosen that they may be broken up into a finite number of n-cells.

3. **Integrals over oriented regions.** To overcome the difficulty mentioned at the end of § 1, we use oriented regions of integration. We define the integral

$$
(10) \qquad \int n \int A_{12\ldots n}\, dx_1 \ldots dx_n,
$$

taken over an oriented n-cell in the space of the x's (of $m > n$ dimensions) given by equations of the form (8) with an order of the u's, as the ordinary integral of § 1 in the space of the u's

$$
(11) \qquad \int n \int \bar{A}_{12\ldots n} \frac{D(x_1 \ldots x_n)}{D(u_1 \ldots u_n)}\, du_1 \ldots du_n.
$$

* For a more complete discussion of the points of this section, see O. Veblen, The Cambridge Colloquium, Analysis situs, p. 100 f.; p. 129 f.

For regions consisting of more than one cell, we evaluate the integral separately over each cell and add the results together.

We may deduce several properties of the integral (10) from this definition. First, the order of the dx's in (10) is an essential factor, since interchanging two of them will change the sign of the Jacobian in (11) and hence change the sign of our result. This also shows that the integral is zero if two of the differentials refer to the same variable or have the same subscript. It also changes sign when a pair of the u's are interchanged, i. e., when the sense of the region of integration is reversed.

The definition preserves two properties of ordinary integrals, one slightly modified. It is distributive with respect to addition, i. e., if $A+B=C$, the same is true of the integrals of these functions, taken over the same region. Also if the variables are changed, the space being referred to a new co-ordinate system, the integral is transformed by the Jacobian, i. e.,

$$(12)\qquad \int^n\int A_{1\ldots n}\,dx_1\cdots dx_n=\int^n\int A_{1\ldots n}\frac{D(x_1\cdots x_n)}{D(y_1\cdots y_n)}\,dy_1\cdots dy_n.$$

4. **Integrals of multilinear forms.** The integral of a multilinear form of the nth order in the m differentials $dx_1\cdots dx_m$,

$$(13)\qquad \int^n\int\sum A_{i_1\ldots i_n}\,dx_{i_1}\cdots dx_{i_n},$$

where the summation is extended over all groups of n, is obtained by summing the results obtained for the separate terms by § 3.

Such an integral can always be replaced by one in which the A's are *skew-symmetric*, i. e., are such that any two A's, with the same subscripts but in different arrangements, have the same numerical value and the same or opposite signs according as the arrangements differ by an even or an odd number of interchanges. To verify this fact, we notice that all the terms involving a given set of differentials $dx_{i_1}\cdots dx_{i_n}$ can be replaced by a single term $B\,dx_{i_1}\cdots dx_{i_n}$ and this in turn by $n!$ terms, one for each arrangement of these m subscripts, the coefficients being $\pm B/n!$ according as their arrangements may be reduced to that of the single term above by an even or an odd permutation of the differentials. Evidently all terms with two or more subscripts identical in (13) may be omitted without changing its value. We thus see that the integral of any form may be reduced to that of a skew-

symmetric form, and in future we shall assume all the forms we are dealing with have been so reduced.

Certain skew-symmetric forms have the property that their integrals taken over any closed n-region, is zero. Consequently their integrals taken over any open n-region, only depend on the $n-1$-region bounding it; and this second condition is completely equivalent to the first statement. Such forms are said to be *integrable*. Our next problem is to determine necessary and sufficient conditions for an integrable form, or *integrability conditions*.

5. **The generalized Green's theorem.** To deduce these conditions for (13), we shall require a generalization of Green's theorem for triple integrals, which we proceed to prove. The theorem is

$$\begin{aligned}(14)\qquad &\int^{n-1}\int \sum A_{i+1\ldots n,1\ldots i-1}\, dx_{i+1}\cdots dx_n\, dx_1 \cdots dx_{i-1}\\ &= \int^{n}\int \sum (-1)^{(n-1)(i-1)} \frac{\partial A_{i+1\ldots n,1\ldots i-1}}{\partial x_i}\, dx_1 \cdots dx_n,\end{aligned}$$

where the first integral is taken over the $(n-1)$-region bounding the n-region over which the second is taken with suitable sense.

We shall prove this equality by evaluating the integral on the right. Let the n-region (in the n-space of the x's) be bounded by the $(n-1)$-region given by the equations

$$(15)\qquad x_i = f_i(u_1 \cdots u_{n-1}), \qquad 0 < i \leq n.$$

Since (15) bounds a portion of n-space, each normal to it (i. e., the straight line perpendicular to the hyperplane at its point of contact) will have an outer and an inner direction. Let us call the outward direction the positively directed normal, and denote its direction cosines by $\alpha_1 \cdots \alpha_n$. If we take an $(n-1)$ cell in (15) associated with a single point of the $(n-1)$ cell, we can then assign a sense to it in such a way that this sensed $(n-1)$ cell, together with the positively directed normal at the point, gives a pre-assigned sense; for example, a sense similar to that of the $x_2 \cdots x_n$ hyperplane (with sense determined by this order of the axes) taken together with the positive x_1-axis. Since the positive sense on the normal is never discontinuous as the normal is varied from point to point, the sense just described will be continuous from point to point.

From (15) we have:

$$dx_i = \sum_j \frac{\partial f_i}{\partial u_j} du_j. \tag{16}$$

By the definition of $\alpha_1 \cdots \alpha_n$

$$\sum \alpha_i \, dx_i = 0; \tag{17}$$

consequently, since the u's are independent,

$$\sum \alpha_i \frac{\partial f_i}{\partial u_j} = 0. \tag{18}$$

These equations show that

$$\alpha_i = K(-1)^{i-1} \frac{D(x_1 \cdots x_{i-1} x_{i+1} \cdots x_n)}{D(u_1 \cdots u_{n-1})}, \tag{19}$$

where K is the same for all α's and is a function determined by the values of the u's. (19) may be re-written

$$\alpha_i = K(-1)^{(n-1)(i-1)} \frac{D(x_{i+1} \cdots x_n x_1 \cdots x_{i-1})}{D(u_1 \cdots u_{n-1})}, \tag{20}$$

since moving the first column into the last place in the numerator of the Jacobian involves $(n-2)$ changes of sign, and this must be done for $(i-1)$ terms, giving for the exponent of (-1),

$$(i-1) + (n-2)(i-1) = (n-1)(i-1).$$

K is a continuous function of the u's throughout the $n-1$ region and is never zero; consequently it preserves its sign. We may make this sign positive by a proper choice of the order of the u's, or sense of the $n-1$ region, and in what follows shall regard K as positive, since this gives the sense required by the theorem.

We can now evaluate the right member of (14). A single term of this integral is

$$\int n \int (-1)^{(n-1)(i-1)} \frac{\partial A_{i+1 \ldots i-1}}{\partial x_i} dx_1 \cdots dx_n. \tag{21}$$

If we perform the integration with respect to x_i, this becomes the ordinary integral

$$(22) \qquad \int^{n-1} \int (-1)^{(n-1)(i-1)} (A' - A'')\, dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n$$

where A' denotes the value of $A_{i+1 \ldots i-1}$ at points where a line parallel to the x_i-axis leaves the n region bounded by (15), i. e., goes from points on the same side of the $n-1$ region as the negatively directed normal to those on the same side of the region as the positively directed normal; while A'' denotes the value of $A_{i+1 \ldots i-1}$ at points where the parallel enters the n-region. For points of the first kind we have

$$(23) \qquad \alpha_i > 0,$$

while for those of the second kind,

$$(24) \qquad \alpha_i < 0.$$

Consequently (cf. 20), at points for which A' is to be evaluated,

$$(25) \qquad \begin{aligned} &|dx_1 \cdots dx_{i-1}\, dx_{i+1} \cdots dx_n| \\ &= (-1)^{(n-1)(i-1)} \frac{D(x_{i+1} \cdots x_n\, x_1 \cdots x_{i-1})}{D(u_1 \cdots\cdots\cdots\cdots u_{n-1})}\, du_1 \cdots du_{n-1}, \end{aligned}$$

while, at points for A'', the absolute value on the left is the negative of this expression. Thus the value of the ordinary integral (22) is equal to that of the integral

$$(26) \qquad \int^{n-1} \int A_{i+1 \ldots i-1}\, dx_{i+1} \cdots dx_{i-1},$$

taken over the bounding $n-1$ region with the sense defined above.

The equivalence of (26) and (21), the first a characteristic term of the left member of (14) and the second a characteristic term of the right member, shows that these members are equal, and thus demonstrates our theorem.

6. Consequences of the theorem. Integrability conditions. The proof just given of the fact that

$$
\begin{aligned}
(27)\qquad & \int n-1 \int A_{i+1\ldots i-1}\, dx_{i+1} \cdots dx_{i-1} \\
& \qquad = \int n \int (-1)^{(n-1)(i-1)} \frac{\partial A_{i+1\ldots i-1}}{\partial x_i}\, dx_1 \cdots dx_n,
\end{aligned}
$$

holds even if the A's involve other variables x_j $(n < j)$ given in terms of the u's by equations of the form (15); consequently (27) is true if the A's are functions of position in an m-space, and the regions of integration are situated in this space.

We may avoid the sign factor in (27), by writing (as suggested to the writer by Professor J. W. Alexander) the differential in the right member in the form $dx_i\, dx_{i+1} \cdots dx_n\, dx_1 \cdots dx_{i-1}$, giving, with new notation

$$
(28)\qquad \int n \int A_{1\ldots n}\, dx_1 \cdots dx_n = \int n+1 \int \frac{\partial A_{1\ldots n}}{\partial x_{n+1}}\, dx_{n+1}\, dx_1 \cdots dx_n.
$$

If we consider the integral (13), by combining equations of the form just written we obtain

$$
\begin{aligned}
(29)\qquad & \int n \int \sum A_{i_1\ldots i_n}\, dx_{i_1} \ldots dx_{i_n} \\
& \qquad = \frac{1}{m-n} \int n+1 \int \sum \frac{\partial A_{i_1\ldots i_n}}{\partial x_{i_{n+1}}}\, dx_{i_{n+1}}\, dx_{i_1} \ldots dx_{i_n},
\end{aligned}
$$

where the summation includes all the permutations of the m subscripts.

This immediately gives the integrability conditions for the integral on the left. For, if it is to be zero for every closed region, the right member must vanish. Consequently the coefficient of any one combination of differentials, as

$$
\sum_j (-1)^{j-1} \frac{\partial A_{i_1\ldots i_{j-1}\, i_{j+1}\ldots i_{n+1}}}{\partial x_{i_j}},
$$

which is the coefficient of $dx_{i_1}\, dx_{i_2} \cdots dx_{i_{n+1}}$, must be zero; and conversely, if all these expressions are zero, the left member will be zero for all regions capable of being the boundary of an $n-1$ region, i. e., for all closed regions.

Accordingly, we have, as the integrability conditions of

$$\int n \int \sum A_{i_1 \ldots i_n} \, dx_{i_1} \cdots dx_{i_n}, \tag{30}$$

the summation involving all the permutations from m subscripts; the $m!/(m-n-1)!(n+1)!$ conditions

$$\sum_j (-1)^{j-1} \frac{\partial A_{i_1 \ldots i_{j-1} i_{j+1} \ldots i_{n+1}}}{\partial x_{i_j}} = 0, \tag{31}$$

or, as they are more usually written,

$$\sum_j (-1)^{n(j-1)} \frac{\partial A_{i_{j+1} \ldots i_{n+1} i_1 \ldots i_{j-1}}}{\partial x_{i_j}} = 0. \tag{32}$$

It is of some interest to write these out in full for the first two orders, as in these cases they are more or less well known. Thus for line-integrals ($n = 1$) equations (32) become of the type

$$\frac{\partial A_i}{\partial x_j} - \frac{\partial A_j}{\partial x_i} = 0 \tag{33}$$

the components of the "curl", while for surface integrals we have

$$\frac{\partial A_{ij}}{\partial x_k} + \frac{\partial A_{jk}}{\partial x_i} + \frac{\partial A_{ki}}{\partial x_j} = 0, \tag{34}$$

the generalization of the curl appearing in the relativistic statement of Maxwell's Equations.*

7. Properties of the integrability conditions. Since the integral of an integrable function over an $n+1$ region only depends on the n region bounding it, we should expect it to be expressible as an integral over this boundary. We may write a given integral

$$\int n+1 \int B_{i_1 \ldots i_{n+1}} \, dx_{i_1} \cdots dx_{i_{n+1}} \tag{35}$$

* Cf. F. D. Murnaghan, The absolute significance of Maxwell's equations, Physical Review, vol. 17 (1921), p. 73.

over any open region as an integral

$$\int^n \int A_{i_1 \ldots i_n} dx_{i_1} \cdots dx_{i_n}, \tag{36}$$

taken over its boundary, by using (29) provided we can find A's satisfying the equations

$$B_{i_1 \ldots i_{n+1}} = \sum (-1)^{j-1} \frac{\partial A_{i_1 \ldots i_{j-1} i_{j+1} \ldots i_{n+1}}}{\partial x_j}. \tag{37}$$

We shall prove that it is necessary and sufficient for the existence of such A's, (or the expression of the B's as "integrability conditions" for an integral of the nth order) that the B's shall satisfy the integrability conditions

$$\sum_j (-1)^{j-1} \frac{\partial B_{i_1 \ldots i_{j-1} i_{j+1} \ldots i_{n+2}}}{\partial x_j} = 0. \tag{38}$$

To prove the conditions necessary, we have merely to notice that, if (37) holds, a term of the form

$$\frac{\partial^2 A_{i_1 \ldots i_{p-1} i_{p+1} \ldots i_{q-1} i_{q+1} \ldots i_{n+1}}}{\partial x_p \, \partial x_q} \tag{39}$$

will appear twice in the sum (38) once from the term in $B_{i_1 \ldots i_{p-1} i_{p+1} \ldots i_{n+1}}$ with coefficient $(-1)^{p+q-3}$, and once from the corresponding term in q with the negative of this as coefficient.

For the sufficiency we first consider the case where we have only $n+2$ subscripts, so that we have $n+2$ equations of the form

$$B_{1 \ldots p-1, p+1 \ldots n+2} = \sum_j (-1)^{j-*} \frac{\partial A_{1 \ldots p-1, p+1 \ldots j-1, j+1 \ldots n+2}}{\partial x_j} \tag{40}$$

to be solved for the A's, and one condition

$$\sum_j (-1)^{j-1} \frac{\partial B_{1 \ldots j-1, j+1 \ldots n+2}}{\partial x_j} = 0. \tag{41}$$

---

* j or $j-1$ as $p <$ or $> j$.

If we write

$$(42)\quad B_{1\ldots p-1,p+1\ldots n+2} = B'_{1\ldots p-1,p+1\ldots n+1} + (-1)^{n-2}\frac{\partial A_{1\ldots p-1,p+1\ldots n+1}}{\partial x_{n+2}},$$

where $p \nmid n+2$, and B' is defined to make this equation agree with the corresponding one of (40); determine the $A_{1\ldots p-1,p+1\ldots n+1}$ to satisfy the remaining equation of (40) (otherwise arbitrary)

$$(43)\qquad B_{1\ldots n+1} = \sum_j (-1)^{j-1}\frac{\partial A_{1\ldots j-1,j+1\ldots n+1}}{\partial x_j},$$

and substitute (42) and (43) in (41); this becomes

$$(44)\qquad \sum_j (-1)^{j-1}\frac{\partial B'_{1\ldots j-1,j+1\ldots n+1}}{\partial x_j} = 0.$$

But, as the B''s are connected with the A's by equations similar to (40) but with one less variable, and (44) bears the same relation to (41), the problem of determining the remaining A's is the same as our original problem for one less variable. Hence we may reduce the problem step by step until we have to determine a single function $A_{3\ldots n+2}$ from the equations

$$\frac{\partial A_{3\ldots n+2}}{\partial x_1} = B''_1, \qquad \frac{\partial A_{3\ldots n+2}}{\partial x_2} = B''_2,$$

with the condition

$$\frac{\partial B''_2}{\partial x_1} - \frac{\partial B''_1}{\partial x_2} = 0,$$

which we know is possible.

For $n+3$ subscripts, we have two additional conditions and $n+2$ additional equations. We first obtain values for the A's not involving the subscript $n+3$ by the method just outlined, using the equations and condition not involving the new subscript. Each of the equations not used will now contain just one A which has been evaluated. As before we write

$$(45)\quad B_{1\ldots p-1,p+1\ldots q-1,q+1\ldots n+3}$$
$$= B'_{1\ldots p-1,p+1\ldots q-1,q+1\ldots n+2} + (-1)^{n+2}\frac{\partial A_{1\ldots p-1,p+1\ldots q-1,q+1\ldots n+2}}{\partial x_{n+3}}$$

to separate out the A's already determined, and, on substituting in the conditions, reduce the problem to that for one order lower, and finally to the case of the first order, which reduces to known conditions.

We may evidently continue the induction to prove the theorem true in general.

It may be noted here that the integrability conditions have one further property, important in tensor analysis. If they are formed from functions A which are the components of a skew-symmetric tensor, they will themselves be tensors. This is easily proved by summing, instead of the derivatives, the covariant derivatives with the proper signs. The additional terms cancel in pairs, owing to the skew-symmetry, and, as the integrability conditions may be represented as sums of covariant derivatives, their tensor character follows from that of the original A's.

8. **Complex integrals of the *m*th order in *m*-space.** An integral of the mth order in complex Euclidean m-space, with co-ordinates $z_1 \cdots z_m$ $(z_k = x_{2k-1} + i x_{2k})$, as

$$\int^m \int F(z_1 \cdots z_m)\, dz_1 \cdots dz_m, \tag{46}$$

where the region of integration is given by a correspondence with a part of the real m-space of the real variables $p_1 \cdots p_m$ is defined as

$$\int^m \int F(z_1 \cdots z_m) \frac{D(z_1 \cdots z_m)}{D(p_1 \cdots p_m)}\, dp_1 \cdots dp_m. \tag{47}$$

This is equivalent to

$$\int^m \int (P + iQ)(dx_1 + i\, dx_2) \cdots (dx_{2k-1} + i\, dx_{2k}) \cdots (dx_{2m-1} + i\, dx_{2m}), \tag{48}$$

where $F = P + iQ$ and (48) is to be evaluated by performing the indicated multiplication (without changing the order of the factors) and interpreting the real integrals resulting by the definition given in § 3.

Confining ourselves to the *real* part of this expansion, we notice that a given set of dx's will occur at most once, and that its coefficient will be $\pm P$ or $\pm Q$. When put in the skew-symmetric form (cf. § 4) the coefficient will be $\pm P/m!$ or $\pm Q/m!$, with a change of sign from the coefficient just mentioned if the order of the factors differs from that in the product of (48) by an odd number of transpositions.

The integrability conditions for the real part of (48) are given by (31):

$$(31) \qquad \sum_j (-1)^{j-1} \frac{\partial A_{i_1 \cdots i_{j-1} i_{j+1} \cdots i_{m+1}}}{\partial x_{i_j}} = 0.$$

If the group of $(m+1)$ i's includes two or more pairs of numbers of the form $(2k-1, 2k)$, all the terms will be identically zero. To give a non-zero result, the $(m+1)$ i's selected must contain only one pair of this form (all groups evidently must contain at least one such pair) and a single integer from each other pair (i. e., it must correspond to a set which contains both differentials from one factor and one differential from each of the other factors of (48)). As a re-arrangement of the i's selected merely changes the sign of all the terms of (31) at most, we may assume without loss of generality that, for the case under consideration, $i_1 = 2k-1$, $i_2 = 2k$, these being the pair of subscripts corresponding to the differentials in a single factor. In this case all the terms of (31) will be zero except

$$(49) \qquad \frac{\partial A_{i_2 \ldots i_{m+1}}}{\partial x_{i_1}} - \frac{\partial A_{i_1 i_3 \ldots i_{m+1}}}{\partial x_{i_2}} = 0.$$

If the first term is $\pm P/m!$, the second is $\pm Q/m!$; while if the first is $\pm Q/m!$, the second is $\mp P/m!$. Thus the last equation reduces either to

$$(50) \qquad \frac{\partial P}{\partial x_{i_1}} - \frac{\partial Q}{\partial x_{i_2}} = 0, \text{ or } \frac{\partial Q}{\partial x_{i_1}} + \frac{\partial P}{\partial x_{i_2}} = 0.$$

These equations are precisely the Cauchy-Riemann differential equations which are necessary and sufficient conditions for the function $F = P + iQ$ being analytic in $z_k = x_{2k-1} + ix_{2k}$. By analogous reasoning we find that the integrability conditions (31) applied to the imaginary part of (48) reduce to equations (50).

9. **Complex integrals of the nth order in m-space.** An integral in a complex m-space of the nth order,

$$(51) \qquad \int n \int F(z_1 \ldots z_m)\, dz_1 \ldots dz_n,$$

taken in a complex m-region given by m equations of the form

$$x_i = f_i(u_1 \ldots u_n), \tag{52}$$

where the u's are complex co-ordinates corresponding to a complex Euclidean n-space, is defined as

$$\int n \int \bar{F}(u_1 \ldots u_n) \frac{D(z_1 \ldots z_n)}{D(u_1 \ldots u_n)} du_1 \ldots du_n, \tag{53}$$

an integral of the type discussed in § 8. With this definition, the integral depends on the order of the differentials, changing sign when two of them are interchanged, and consequently the procedure of § 4 will enable us to write (51), or the integral of any multilinear form in the complex differentials, as the integral of a skew-symmetric form. However, this is unnecessary for the application of the integrability conditions, as if the functions F and the f_i are analytic, the argument of § 8 shows that these conditions will be satisfied identically. The theorem that (51) is integrable provided F and the f_i are analytic functions, is a generalization of Cauchy's theorem on the integral of a function of a single complex variable.

HARVARD UNIVERSITY.

ON SYMMETRIC FORMS IN N VARIABLES.

By Arnold Dresden.

The theory of symmetric functions as discussed in the treatises on Algebra* contains a proof of the theorem that every polynomial symmetric in the variables $a_1, a_2, \ldots a_n$, can be expressed as a polynomial in the elementary symmetric functions of these variables, i. e., in terms of $\sum a_1, \sum a_1 a_2, \ldots . \sum a_1 a_2 \cdots a_{n-1}, a_1 a_2 \cdots a_n$. There is moreover a quite extensive literature devoted to the problem of calculating the latter polynomials. A historical survey of this question is found in the Introduction to Decker's Tables of the Fifteenthic,† to which the reader is referred for an account of the work done on this subject. This work has made possible a considerable simplification in the calculations necessary for tables of symmetric functions; but it does not make the use of such tables indispensable. As such tables can not for a long time become available beyond a limited range as to weight, it has seemed to the writer desirable to return to the problem of establishing a general formula which shall enable one to express a symmetric form of arbitrary weight, arbitrary number of parts and arbitrary distribution of the total weight over the parts, in terms of the elementary symmetric functions. To this question the present paper is devoted. We shall use throughout the symbolic notation‡ in which only the exponents of the factors in the representative term of the form are written and in which repetition of the same exponent is indicated by an upper index; thus the forms $\sum a_1^4 a_2^4 a_3^2 a_4 a_5$ and $\sum a_1^5 a_2^5 a_3^4 a_4^3 a_5^3 a_6^3 a_7^2 a_8$ are represented by the symbols $(4^2 21^2)$ and $(5^2 43^3 21)$ respectively.

1. We begin by considering products of the form $(1^{s_1})(1^{s_2}) \cdots (1^{s_n})$, $s_1 \geqq s_2 \geqq \cdots \geqq s_n$, where, in accordance with the notation explained above, (1^{s_i}) designates the symmetric form of s_i parts, each of weight 1, i. e., the elementary symmetric function $\sum a_1 a_2 \cdots a_{s_i}$.

The product $(1^{s_1})(1^{s_2})$ will consist of terms of the type $\sum a_1^2 \cdots a_{s_2-j}^2 a_{s_2-j+1} \cdots a_{s_1+j}$, $(j = 0, 1, \ldots, s_2)$. A term of this type is obtained whenever j factors of (1^{s_2}) are used to form some of the $s_1 - s_2 + 2j$ first powers in the product.

* As e. g. Serret, Cours de'Algèbre, vol. 1.

† F. F. Decker, The symmetric function tables of the fiftheenthic, published by the Carnegie Institution of Washington, 1910.

‡ See e. g. MacMahon, Combinatory analysis, p. 1—7.

Hence it can be obtained in $\binom{s_1-s_2+2j}{j}$ ways and will consequently appear in the product with this coefficient. Therefore we shall have

$$(1)\qquad (1^{s_1})(1^{s_2}) = \sum_{j=0}^{s_2}\binom{s_1-s_2+2j}{j}(2^{s_2-j}\,1^{s_1-s_2+2j}).$$

As a check on this result we count the number of terms on each side of this equation. On the left there are $\binom{n}{s_1}\binom{n}{s_2}$, while the number of terms on the right is $\sum_{j=0}^{s_2}\binom{s_1-s_2+2j}{j}\binom{n}{s_2-j}\binom{n-s_2+j}{s_1-s_2+2j} = \binom{n}{s_1}\sum_{j=0}^{s_2}\binom{s_1}{s_2-j}\binom{n-s_1}{j}$. Since now $\binom{n}{s_2} = \sum_{j=0}^{s_2}\binom{s_1}{s_2-j}\binom{n-s_1}{j}$*, it follows that the number of terms on the two sides of equation (1) are equal to each other.

If both sides of equation (1) be multiplied by (1^{s_3}), $s_3 < s_2$, and the result simplified, we obtain the equation

$$(1^{s_1})(1^{s_2})(1^{s_3}) = \sum\sum\binom{s_1-s_2+2k_{21}}{k_{21}}\binom{s_2-s_3+2k_{32}+k_{31}-k_{21}}{k_{32}}$$

$$\binom{s_1-s_2+2k_{21}+k_{31}-k_{32}}{k_{31}}(3^{s_3-k_{32}-k_{31}}\,2^{s_2-s_3+2k_{32}+k_{31}-k_{21}}\,1^{s_1-s_2+2k_{21}+k_{31}-k_{32}}),$$

where the first sum is to be taken with respect to k_{21} from 0 to s_2 and the second sum with respect to k_{31} and k_{32} from 0 to s_3, under the restriction $k_{31}+k_{32} \leq s_3$. Besides the indices k_{21}, k_{31} and k_{32}, we now introduce the indices k_{11}, k_{22} and k_{33}, such that $k_{11} = s_1$, $k_{21}+k_{22} = s_2$ and $k_{31}+k_{32}+k_{33} = s_3$. If, moreover, we write $\sum_{m=j}^{i} k_{mj} = \sigma_{ij}$, for $j \leqq i$, while $\sigma_{ij} = 0$, for $j > i$, we can write the above result in the following form

$$(1^{s_1})(1^{s_2})(1^{s_3}) = \sum\sum\binom{\sigma_{21}-\sigma_{22}}{k_{21}}\binom{\sigma_{31}-\sigma_{32}}{k_{31}}\binom{\sigma_{32}-\sigma_{33}}{k_{32}}(3^{\sigma_{31}}\,2^{\sigma_{22}-\sigma_{33}}\,1^{\sigma_{11}-\sigma_{22}}),$$

in which the sums are to be carried out as specified above.

* See Netto, Combinatorik, p. 13, or Chrystal, Algebra, vol. 2, p. 8.

This formula leads us to conjecture the following general formula

$$(2)\qquad \prod_{l=1}^{n}(1^{s_l}) = \sum \prod_{i=1}^{n}\prod_{j=1}^{i}\binom{\sigma_{ij}-\sigma_{i,j+1}}{k_{ij}}\left(\prod_{l=1}^{n} l^{\sigma_{n,l}-\sigma_{n,l+1}}\right),\quad s_1 \geqq s_2 \geqq \cdots \geqq s_n,$$

in which $\sum$ designates an n-fold sum, the i-th partial sum of which is to be taken over all i-partite partitions of s_i, consisting of positive or zero integers $k_{i1}, \cdots, k_{ii}$ and designated by (k_{ij}), and where

$$(3)\qquad \sigma_{ij} = \sum_{m=j}^{i} k_{mj} \text{ for } j \leqq i, \text{ and } \sigma_{ij} = 0 \text{ for } j > i.$$

Since formula (2) is evidently satisfied for $n = 2, 3$, its general validity will be established as soon as we have proved that it holds for $n+1$, assuming that it holds for n.

2. For this purpose, we consider first the product $(1^{s_{n+1}})\left(\prod_{l=1}^{n} l^{\sigma_{n,l}-\sigma_{n,l+1}}\right)$. If $k_{n+1,j}$ $(j = 1, \cdots, n+1)$ be an arbitrary $(n+1)$-partite partition of s_{n+1}, consisting of positive or zero elements, let $k_{n+1,l}$ of the s_{n+1} factors in $(1^{s_{n+1}})$ be used to form l-th powers in the product out of $(l-1)$th powers in its second factor. Of the $\sigma_{n,l}-\sigma_{n,l+1}$ l-th powers in this second factor, $k_{n+1,l+1}$ will then be used in the $(l+1)$th powers in the product, so that the total number of factors of weight l in the product will be $k_{n+1,l}+\sigma_{n,l}-\sigma_{n,l+1}-k_{n+1,l+1} = \sigma_{n+1,l}-\sigma_{n+1,l+1}$. Hence the product will consist of terms of the form $\left(\prod_{l=1}^{n+1} l^{\sigma_{n+1,l}-\sigma_{n+1,l+1}}\right)$. But the part of weight l in this product may be obtained by using for the new l-th powers any $k_{n+1,l}$ out of $\sigma_{n+1,l}-\sigma_{n+1,l+1}$ factors, i. e. in $\binom{\sigma_{n+1,l}-\sigma_{n+1,l+1}}{k_{n+1,l}}$ ways. We conclude that the coefficient of the general term in the product will be $\prod_{l=1}^{n+1}\binom{\sigma_{n+1,l}-\sigma_{n+1,l+1}}{k_{n+1,l}}$. Consequently we find that

$$(4)\qquad (1^{s_{n+1}})\left(\prod_{l=1}^{n} l^{\sigma_{n,l}-\sigma_{n,l+1}}\right) = \sum \prod_{l=1}^{n+1}\binom{\sigma_{n+1,l}-\sigma_{n+1,l+1}}{k_{n+1,l}}\left(\prod_{l=1}^{n+1} l^{\sigma_{n+1,l}-\sigma_{n+1,l+1}}\right).$$

in which the sum is to be taken over all $(n+1)$-partite partitions $(k_{n+1,l})$ of s_{n+1}. If we now multiply both sides of (2) by $(1^{s_{n+1}})$ and then make use of (4), we obtain the following result:

$$\prod_{l=1}^{n+1}(1^{s_l}) = \sum \prod_{i=1}^{n+1} \prod_{j=1}^{i} \binom{\sigma_{ij}-\sigma_{i,j+1}}{k_{ij}} \left(\prod_{l=1}^{n+1} l^{\sigma_{n+1,l}-\sigma_{n+1,l+1}}\right),$$

in which $\sum$ now represents an $(n+1)$-fold sum, the i-th partial sum of which is to be taken over all i-partite partitions of s_i. But this is exactly what would result from (2) if $n+1$ were substituted for n. Hence the validity of (2) has been established.

3. If in the sum on the right of (2) we put $k_{ij} = 0$ for $j < i$ and $k_{ii} = s_i$, all the factors of the coefficient $\prod_{i=1}^{n}\prod_{j=1}^{i}\binom{\sigma_{ij}-\sigma_{i,j+1}}{k_{ij}}$ for which $j < i$, are obviously equal to 1; while, for $j = i$, we have $\sigma_{ij} = \sigma_{ii} = k_{ii}$, and $\sigma_{i,j+1} = 0$, so that the corresponding factors in the coefficient are also equal to 1. Moreover, $\sigma_{n,l}$ reduces in this case to k_{ll}, which is equal to s_l, so that we obtain for those values of k_{ij} the term $\left(\prod_{l=1}^{n} l^{s_l - s_{l+1}}\right)$, which is precisely the general symmetric form with which we are concerned. We shall call this term the "leading term" in the sum on the right of (2). We now wish to set up a system of equations analogous to (2) and such that between them and (2) we can eliminate all but the leading term of (2).

From the definition of σ_{ij} in (3), it follows that $\sum_{l=1}^{n}\sigma_{nl} = \sum_{l=1}^{n}\sum_{m=l}^{n}k_{ml} = \sum_{m=1}^{n}\sum_{l=1}^{m}k_{ml} = \sum_{m=1}^{n}s_m$, so that if we write $\sum_{m=1}^{n}s_m = S$, the set (σ_{nl}) is an ordered partition of S. But, since the indices k_{ij} are all non-negative integers, it also follows from (3), that $\sum_{l=i}^{n}\sigma_{nl} < \sum_{l=i}^{n}s_l$ and hence, in view of the relation $\sum_{l=1}^{n}\sigma_{nl} = \sum_{l=1}^{n}s_l$, that $\sum_{l=1}^{i}\sigma_{nl} \geq \sum_{l=1}^{i}s_l$. This fact may be expressed by saying that, if generated by increasing the index i from 1 to n, the partition (σ_{nl}) dominates the partition (s_l), while, if generated by decreasing the index i from n to 1, (s_l) dominates (σ_{nl}); we shall now lay down the following definition:

Definition. Two ordered n-partite partitions of S, denoted by $(r_{n,l})$ and $(u_{n,l})$ are related by the asymmetric dominance relation D, $(r_{n,l})\,D\,(u_{n,l})$, provided $\sum_{l=1}^{i}r_{n,l} \geq \sum_{l=1}^{i}u_{n,l}$, for $i = 1, 2, \cdots, n$. Hence we can say that $(\sigma_{nl})\,D\,(s_l)$, where (s_l) is understood to designate the n-partite partition of S, consisting

of the numbers $s_1, \ldots, s_n$, so that every term on the right of (2) has *corresponding to it a partition* (t_l) *such that* $(t_l)\, D\, (s_l)$.* It follows from the nature of the coefficients that only such terms will actually occur in (2) as correspond to ordered partitions (t_l) in which $t_l \geqq t_{l+1}$. Such partitions we shall call non-increasing ordered partitions.

But it is also true that corresponding to every partition (t_l) of S, such that $(t_l)\, D\, (s_l)$, there exists at least one term in the sum on the right of (2). To prove this, we have to show that, if $\sum_{l=1}^{n} t_l = \sum_{l=1}^{n} s_l$, and $\sum_{l=1}^{i} t_l \geqq \sum_{l=1}^{i} s_l$, there exists at least one solution of the system of equations

$$\sum_{j=1}^{i} k_{ij} = s_i, \qquad i = 1, 2, \ldots, n; \tag{5}$$

$$\sum_{i=j}^{n} k_{ij} = t_j, \qquad j = 1, 2, \ldots, n. \tag{6}$$

Equation (5), for $i = 1$ gives us $k_{11} = s_1$. We determine now an integer k_{21} subject to the conditions

$$0 \leqq k_{21} \leqq t_1 - k_{11},$$

and

$$s_2 - t_2 \leqq k_{21} \leqq s_2.$$

From our hypothesis, $(t_l)\, D(s_l)$, it follows that $t_1 - k_{11} = t_1 - s_1 \geqq s_2 - t_2$, so that these two conditions are compatible; there exists therefore at least one non-negative integer, which satisfies these conditions. If we now take $k_{22} = s_2 - k_{21}$, it follows that $0 \leqq k_{22} \leqq t_2$, so that we have found non-negative integers k_{11}, k_{21} and k_{22}, such that $k_{11} = s_1$, $k_{21} + k_{22} = s_2$; $k_{11} + k_{21} \leqq t_1$, and $k_{22} \leqq t_2$.

Next we determine two non-negative integers k_{31} and k_{32} such that

$$0 \leqq k_{31} \leqq t_1 - (k_{11} + k_{21}),$$

$$0 \leqq k_{32} \leqq t_2 - k_{22},$$

$$s_3 - t_3 \leqq k_{31} + k_{32} \leqq s_3.$$

* Since the partitions used hereafter are all n-partite, we shall omit the subscript n from the partition symbol, i. e. we shall write (t_l) in place of (t_{nl}).

The compatibility of these conditions follows again from the fact that $(t_l)D(s_l)$; for this implies that $s_3 - t_3 \leqq t_1 + t_2 - (s_1 + s_2) = t_1 + t_2 - (k_{11} + k_{21} + k_{22})$. We now define $k_{33} = s_3 - (k_{31} + k_{32})$ and complete in this way the determination of the non-negative integers k_{11}, k_{21}, k_{22} and k_{31}, k_{32}, k_{33}, such that $k_{11} = s_1$, $k_{21} + k_{22} = s_2$, $k_{31} + k_{32} + k_{33} = s_3$, while $k_{11} + k_{21} + k_{31} \leqq t_1$, $k_{22} + k_{32} \leqq t_2$, $k_{33} \leqq t_3$.

Suppose now that we have determined $i-1$ sets of non-negative integers $k_{l1}, \ldots, k_{ll}$, $l = 1, 2, \ldots, i-1$, such that

$$\sum_{j=1}^{l} k_{lj} = s_l, \tag{5a}$$

and

$$\sum_{j=l}^{i-1} k_{jl} \leqq t_l. \tag{6a}$$

We determine then $i-1$ new integers $k_{i1}, \ldots, k_{i,i-1}$ so as to satisfy the conditions

$$0 \leqq k_{ij} \leqq t_j - \sum_{l=j}^{i-1} k_{lj},\ j = 1, \ldots, i-1,$$

and

$$s_i - t_i \leqq \sum_{j=1}^{i-1} k_{ij} \leqq s_i.$$

From the dominance hypothesis, we derive now the fact that $s_i - t_i \leqq \sum_{j=1}^{i-1} t_j - \sum_{j=1}^{i-1} s_j = \sum_{j=1}^{i-1} t_j - \sum_{j=1}^{i-1} \sum_{l=j}^{i-1} k_{lj}$, so that there exists at least one set of integers $k_{i1}, \ldots, k_{i,i-1}$ satisfying these conditions. If we now set $k_{ii} = s_i - \sum_{j=1}^{i-1} k_{ij}$, we shall have determined i-sets of non-negative integers which satisfy equations (5a) and (6a), in which $i-1$ has been replaced by i. Since the existence of such sets has already been shown for $i = 1, 2, 3$, we conclude that there exists at least one solution for equations (5a) and (6a) for $i = n$. To complete the solution of equations (5) and (6) we set

$$k_{nj} = t_j - \sum_{l=j}^{n-1} k_{lj} \text{ for } j = 1, \ldots, n-1,$$

and

$$k_{nn} = t_n.$$

It follows then from (6a), written for $i = n$, that $k_{nj} \geq 0, j = 1, \ldots, n-1$; and, from (5a), written for $i = n$, we conclude, that $\sum_{j=1}^{n} k_{nj} = \sum_{j=1}^{n} t_j - \sum_{j=1}^{n-1} \sum_{l=j}^{n-1} k_{lj}$ $= \sum_{j=1}^{n} t_j - \sum_{j=1}^{n-1} s_j = s_n$.

We may now conclude that the sum on the right of (2) contains all those terms which correspond to non-increasing ordered partitions (t_l), for which $(t_l)\ D(s_l)$, and no others. We see furthermore that the dominance relation D is transitive in the field of non-increasing ordered partitions, so that if we write a formula analogous to (2), but replacing (s_l) by (t_l), where $(t_l)\ D(s_l)$, then the terms on the right-hand side of this new formula will also correspond to non-increasing ordered partitions of S which bear to (s_l) the relation D.

4. We now write for every non-increasing partition (t_l) for which the relation $(t_l)\ D(s_l)$ holds, an equation analogous to (2) and obtained from (2) by replacing (s_l) by (t_l).* In virtue of the remark made at the end of 3, the right sides of these equations will contain only such terms as already occur on the right of (2). Since there will be moreover one equation for each term on the right of (2), we shall have a system of linear non-homogeneous equations with these terms as unknowns. But the "leading terms" in these equations have coefficients equal to 1. The determinant of the system will therefore, if properly arranged, have a determinant equal to 1, so that we can obtain from this system of equations the leading term of (2), which is the general symmetric form which we are concerned with, viz. $\left(\prod_{l=1}^{n} l^{s_l - s_{l+1}}\right)$. In order to carry through the determination of this term, we seek multipliers $\lambda(t_1, \ldots, t_n; s_1, \ldots, s_n)$ such that in the equation obtained by adding together the equations written for the partitions (t_l), each of these having been multiplied by $\lambda(t_1, \ldots, t_n; s_1, \ldots, s_n)$, the right side will consist of the single term $\left(\prod_{l=1}^{n} l^{s_l - s_{l+1}}\right)$. This final equation will then give us the desired result, viz.

* All the partitions of S used from now on will be non-increasing ordered partitions. We shall therefore omit these qualifying adjectives in the sequel.

a formula of the following form

$$(7)\qquad \left(\prod_{l=1}^{n} l^{s_l - s_{l-1}}\right) = \sum \lambda(t_1, \ldots, t_n; s_1, \ldots, s_n) \prod_{l=1}^{n} (1^{t_l}).$$

We have therefore obtained the following theorem:

THEOREM. An arbitrary symmetric form $\left(\prod_{l=1}^{n} l^{s_l - s_{l+1}}\right)$ is expressible as a linear function of such products of elementary symmetric functions $\prod_{l=1}^{n}(1^{t_l})$ as are determined by indices t_l such that $(t_l)\ D(s_l)$.

We turn next to the problem of determining the multipliers λ, and we observe at once that we must have $\lambda(s_1, \ldots, s_n; s_1, \ldots, s_n) = 1$.

5. A form $\left(\prod_{l=1}^{n} l^{r_l - r_{l+1}}\right)$ will appear in the formula for $\prod_{l=1}^{n}(1^{t_l})$, analogous to formula (2), provided $(r_l)\ D(t_l)$; if it appears, it will have the coefficient $\sum \prod_{i=1}^{n} \prod_{j=1}^{i} \binom{\tau_{ij} - \tau_{i,j-1}}{k_{ij}}$, the sum being extended over all those sets of indices k_{ij}, which satisfy the equations $\sum_{j=1}^{i} k_{ij} = t_i$ and $\sum_{i=j}^{n} k_{ij} = r_j$, and where $\tau_{ij} = \sum_{m=j}^{i} k_{mj}$. In order to eliminate this form from our final result, we multiply each of the equations in which it appears by a multiplier $\lambda(t_1, \ldots, t_n; s_1, \ldots, s_n)$, chosen in such a way that when these equations are added after having been multiplied through by the proper multiplier, the total coefficient of the form $\left(\prod_{l=1}^{n} l^{r_l - r_{l+1}}\right)$ shall vanish. Hence we obtain for the multiplers λ the following reduction formula

$$(8)\qquad \sum \lambda(t_1, \ldots, t_n; s_1, \ldots, s_n) \sum \prod_{i=1}^{n} \prod_{j=1}^{i} \binom{\tau_{ij} - \tau_{i,j-1}}{k_{ij}} = 0.$$

Here, the outer sum is to be extended over all those partitions (t_l) of S, for which $(r_l)\, D(t_l)\, D(s_l)$, while the inner sum is to be taken for all those sets of indices k_{ij}, which satisfy the equations

(9) $$\sum_{j=1}^{i} k_{ij} = t_i, \text{ and } \sum_{i=j}^{n} k_{ij} = r_j,$$

and

(10) $$\tau_{ij} = \sum_{m=j}^{i} k_{mj}.$$

If we take for (r_l) successively sets which will yield for the outer sum in (8) 2, 3, ... terms etc., this formula (8) will enable us to determine the multipliers λ explicitly.

6. Examples.

To calculate $\sum x_1^2 x_2 x_3 x_4$. This form is represented by the symbol (21^3), so that we have $s_2 = 1$, $s_1 = 4$. The partitions of 5, which bear to (4, 1) the relation D are (4, 1) and (5, 0). We have therefore, by the use of formula (7)

$$(21^3) = \lambda(4, 1; 4, 1)\cdot(1^4)(1) + \lambda(5, 0; 4, 1)\cdot(1^5).$$

Since $\lambda(4, 1; 4, 1) = 1$, it remains to calculate $\lambda(5, 0; 4, 1)$. For this purpose, we make use of formula (8), with $r_1 = 5$ and $r_2 = 0$. There are two partitions (t_l), viz., $t_1 = 5$, $t_2 = 0$; $t_1 = 4$, $t_2 = 1$. The equations for k_{ij} have a unique solution in each of these cases, viz. $k_{11} = 5$, $k_{21} = k_{22} = 0$; and $k_{11} = 4$, $k_{21} = 1$, $k_{22} = 0$. It is now a simple matter to determine the coefficients in equation (8) and we find $\lambda(5, 0; 4, 1) + 5\lambda(4, 1; 4, 1) = 0$, whence $\lambda(5, 0; 4, 1) = -5$. We obtain therefore the following result: $\sum x_1^2 x_2 x_3 x_4 = E_4 E_1 - 5E_5$, where E_i designates the elementary symmetric function of weight i.

To calculate $\sum x_1^2 x_2^2 x_3^2 x_4 x_5$. This form is represented by the symbol $(2^3 1^2)$, so that $s_1 = 5$, $s_2 = 3$. By use of (7) we find

$$(2^3 1^2) = \lambda(5, 3; 5, 3)(1^5)(1^3) + \lambda(6, 2; 5, 3)(1^6)(1^2)$$
$$+ \lambda(7, 1; 5, 3)(1^7)(1) + \lambda(8, 0; 5, 3)(1^8).$$

To determine $\lambda(6, 2; 5, 3)$ we use (8) with $r_1 = 6$ and $r_2 = 2$. We find $\lambda(6, 2; 5, 3) = -4$. Now we use (8) again with $r_1 = 7$ and $r_2 = 1$, and

we find $\lambda(7, 1; 5, 3) = 9$. Using (8) a third time with $r_1 = 8$, $r_2 = 0$, we find $\lambda(8, 0; 5, 3) = -16$. We conclude

$$\sum x_1^2 x_2^2 x_3^2 x_4 x_5 = E_5 E_3 - 4 E_6 E_2 + 9 E_7 E_1 - 16 E_8.$$

When the weight of the form increases and particularly when the number of parts of different weight increases, the labor of the calculation increases a great deal. It becomes therefore desirable to develop general methods for the determination of the multipliers λ, which will enable us to determine them more readily. It is my purpose to develop such methods in a later paper.

THE UNIVERSITY OF WISCONSIN,
August, 1922.

ALGEBRAIC FIELDS.

By J. H. M. Wedderburn.

1. **Introduction.** The object of this paper is purely expository. It is based mainly on the ideas sketched by Kronecker in the first part of his paper "Über den Zahlbegriff"* which in turn have their origin in Cauchy's "Mémoire sur une nouvelle théorie des imaginaires et sur les racines symbolique des équations et des equivalences".†

The trend of the discussion will be best understood by indicating briefly the difference between algebraic and analytical methods. Both algebra and analysis deal with a class of entities — for instance real or complex numbers — subject to two operations addition and multiplication. Algebra is the study of these two relations, generally between a finite number of elements but not necessarily so; analysis on the other hand considers in addition a third relation, namely that of *order* which need not subsist as regards all the elements used but is nevertheless fundamental.

In analysis — or perhaps one should say analytical algebra — a fundamental theorem is that every equation has a root, and the proof of this theorem depends on continuity considerations, and therefore on the idea of order. But a part of algebra at least deals with situations in which there is no question of continuity, as in the theory of finite Galois fields for instance. One should therefore suspect that this theorem should in reality be in no way "fundamental" in algebra proper, and this is indeed the case, as its place may be taken by a theorem to the effect that, if an equation does not have a root in the class of entities under consideration (for instance rational numbers), then that class can be so extended that the given equation does have a root.

The operations $+$ and $\times$ were originally defined for positive integers but it is readily seen that relations with similar properties may subsist in other sets. To take a useful, but rather trivial, example suppose that the set under consideration consists of only two elements a and b. If we define the operations $+$ and $\times$ by

$$a+a=a, \quad a+b=b=b+a, \quad b+b=a, \quad a\times a=a,$$

$$a\times b=a=b\times a, \quad b\times b=b,$$

* Crelle, Bd. 101, pp. 337–355.

† Comptes Rendus, 24 (1847), pp. 1120; Œuvres (1) 10, pp. 312–323.

it is readily shown that all the processes used in algebra but not involving order* are applicable to this set if a is identified with 0. We shall therefore not assume the set of integers as our starting point, as Kronecker does, but shall start with a set of undefined elements subject to two operations obeying certain postulates.† This set is then enlarged when necessary, that is, when a problem arises which leads to an algebraic equation which has no solution in the original set.

There are three principal methods of doing this. In the first place we may add arbitrarily new elements or symbols which satisfy certain conditions as for instance $i^2 = -1$. The logical basis of this method is not usually so secure as in the two that follow. The second method, due to Hamilton,‡ constructs an entirely new set of elements by considering ordered sets of elements in the original system. For these new entities two operations, say $\oplus$ and $\otimes$, are defined so as to satisfy conditions similar to those imposed on $+$ and $\times$. For instance, taking the set B of pairs of elements (a, b) in the original set A, the defining relations for ordinary complex numbers are

1) $(a, b) = (c, d)$ if, and only if, $a = c$, $b = d$.

2) $(a, b) \oplus (c, d) = (a + c, b + d)$.

3) $(a, b) \otimes (c, d) = (ac - bd, ad + bc)$.

When this is done, it is found that B contains a set C simply isomorphic with the set A, namely the set of elements of the form $(a, 0)$. The set A is then superfluous, C taking its place, and so there is no need to maintain any special symbol for $\oplus$ and $\otimes$ which are replaced by $+$ and $\times$ without chance of confusion.

Cauchy's method is quite different. Starting with, say, the set of real numbers, we consider all real polynomials in a real *variable* i. Any such function can be put in the form

$$f(i) = q(i)(i^2 + 1) + ai + b$$

* This does not include arithmetical theorems such as theorems regarding primes and factorization. The set given here is of course the set of integers reduced modulo 2 i.e., $GF[2]$.

† This is in no way meant to controvert Kronecker's thesis that all analysis is based on the set of integers. The sole object is to avoid having to make a fresh start when it is desired to apply operations analogous to addition and multiplication to a set which may perhaps be defined in terms of the set of integers but does not include that set.

‡ See the interesting introduction to his "Lectures on quaternions", Dublin, 1853, where other references are given.

by means of the division transformation. Polynomials are then classified according to the value of the remainder $ai+b$, all those that give the same remainder being said to belong to the same class or to be congruent modulo (i^2+1). As any class is completely determined by the sole linear function it contains, it is convenient to take this linear function as representing its class. Addition $\oplus$ is then defined as ordinary addition and multiplication $\otimes$ as the result of performing ordinary multiplication and retaining only the remainder obtained on dividing by i^2+1. For example, the ordinary product of $ai+b$ and $ci+d$ is $aci^2+(bc+ad)i+bd$; we therefore define $(ai+b)\otimes(ci+d)$ as $(bc+ad)i+bd-ac$, which is the remainder on dividing by i^2+1. The operations are then identified with $+$ and $\times$ as before.

As in the second method, the new set is seen to contain a subset simply isomorphic with the old set, which can therefore be discarded. This is the abstract point of view, but Kronecker emphasizes the fact that every relation in the new set corresponds to an identity in the old; for instance

$$(ai\oplus b)\otimes(ci\oplus b)=(ad+bc)i\oplus(bd-ac)$$

corresponds to the identity

$$(ai+b)(ci+d)=(ad+bc)i+(bd-ac)+ac(i^2+1).$$

Since

$$(x^2+1)=(x-i)(x+i)+(i^2+1)\equiv(x-i)(x+i)\ \text{mod.}\ (i^2+1).$$

the polynomial x^2+1 is reducible in the new set.

Any irreducible function may take the place of i^2+1; for example, α^2-2 is irreducible in the set of rational numbers and, if as above we take the set of rational polynomials in α reduced modulo α^2-2, the function α in this set has the properties of $\sqrt{2}$ since

$$\alpha^2=2+(\alpha^2-2)\equiv 2\ \text{mod.}\ (\alpha^2-2).$$

Kronecker extends this process still farther and, starting with the set of positive integers (including 0), introduces negative and rational numbers by this method. Thus, in place of the arithmetical relation

$$5-9=7-11=-4.$$

he considers the identity

$$5+9x+2(x+1) = 7+11x = 4x+7(x+1),$$

which may be written

$$5+9x \equiv 7+11x \equiv 4x \bmod.(x+1).$$

This is the point of view which we now proceed to develop more fully.*

2. **The fundamental postulates.** For the reasons indicated above instead of taking the positive integers as the basis of our number-system, we shall assume that we are dealing with a set D of elements $a, b, c, \ldots$, finite or infinite in number, which satisfy certain conditions or postulates specified below. No attempt has been made to render these postulates independent. They have been framed to show the similarity that exists between the operations of addition and multiplication, and redundant postulates have been included so as to emphasize this similarity.

It would be out of place here to discuss what is meant by an 'operation'; it is sufficient for our purposes to say that if a and b are any elements of the set, not necessarily distinct, there exist two single-valued functions of a and b, denoted by $a+b$ and $a\times b$ respectively, and that these functions have the following properties:

A 1. $a+b$ *is an element of the given set.*
2. $a+b = b+a$.
3. $a+(b+c) = (a+b)+c$.
4. *If* $a+b = a+c$, *then* $b = c$.
5. *There is an element* 0 *such that* $0+0 = 0$.

M 1. $a\times b$ *is an element of the given set.*
2. $a\times b = b\times a$.
3. $a\times(b\times c) = (a\times b)\times c$.
4. *If* $a\times b = a\times c$ *but* $a+b \neq b$, *then* $b = c$.
5. *There is an element* $1 \neq 0$ *such that* $1\times 1 = 1$.

AM 1. $(a+b)\times c = a\times c+b\times c$.
1′. $c\times(a+b) = c\times a+c\times b$.
2. *If* $a\times c+b\times d = a\times d+b\times c$, *and* $a \neq b$, *then* $c = d$.

A set D which satisfies these postulates will be called a *semi-field*. As an example we may take the set of positive numbers and zero the operations +

* For a more detailed account of the theory of fields, the reader is referred to E. Steinitz, Algebraische Theorie der Körper, Journ. für Math., vol. 137 (1909), pp. 167–309, and J. König, Algebraische Größen, Leipzig, 1903.

and $\times$ being ordinary addition and multiplication; this example shows that division and subtraction are not necessarily possible in a semi-field.

We now state some obvious consequences of these postulates. The postulates used are stated in each case but the details of the proof are in many cases left to the reader.

THEOREM 2.1. AM 1′ *follows from* A 1, M 2, *and* AM 1.

THEOREM 2.2. *If a is any element of the set D, then* $a + 0 = a$.

For

$$0 + (a + 0) = (0 + 0) + a \qquad \text{(A 2, 3)}$$

$$= 0 + a, \qquad \text{(A 5)}$$

and therefore by A 4 we have $a + 0 = a$.

Any element z such that $a + z = a$ for every a is called a *null* element.

THEOREM 2.3. 0 *is the only null element.* (A 2, 3, 4, 5)

For, if $a + z = a$, then, by Theorem 2.2, $a + z = a + 0$ and hence $z = 0$ by A 4.

THEOREM 2.4. $a \times 1 = a$. (M 2, 3, 4, 5)

An element i such that $a \times i = a$ for every element a of the set which is not a null element is called an *identity* element.

THEOREM 2.5. 1 *is the only identity in the set.* (M 2, 3, 4, 5)

The proofs of these two theorems are left to the reader as they are the same as the proofs of 2.2 and 2.3 with $\times$ substituted for $+$.

THEOREM 2.6. *If a is any element of the set, then* $a \times 0 = 0$.
(A 1, 2, 3, 4, 5; AM 1)

For, if b is any element, then

$$a \times b + a \times 0 = a \times (b + 0) \qquad \text{(AM 1)}$$

$$= a \times b \qquad \text{(Th. 2.2)}$$

therefore

$$a \times 0 = 0. \qquad \text{(Th. 2.3)}$$

THEOREM 2.7. M 4 *is implied by* A 1, 2, 3, 4, 5 *and* AM 1, 2.

For, if

$$a \times b = a \times c,$$

then

$$a \times b + 0 \times c = a \times c + 0 \times b. \qquad \text{(Th. 2.6)}$$

Hence it follows from AM 2 that $b = c$ provided $a \neq 0$.

In place of A 5 and AM 2 we might have used the postulate:

A 5′. *If a and b are elements of the set, there is in the set an element c such that either* $a+c=b$ or $a=b+c$.

If we put $a=b=0$ in A 5′, we get A 5 so that the former is considerably stronger as a postulate. In fact, together with A 2, 3, 4 and M 1, 2, 4 it includes AM 2. For, if $a\times c+b\times d=a\times d+b\times c$, $a\neq b$, then there is an element x such that either $a=b+x$ or $b=a+x$ and, owing to the symmetry of AM 2, we may assume without loss of generality that $a=b+x$, $x\neq 0$. We have then

$$b\times c+x\times c+b\times d=b\times d+x\times d+b\times c;$$

therefore

$$x\times c=x\times d,$$

and hence, by M 4, $c=d$.

In the remainder of the paper we shall, as is usual, employ ab in place of $a\times b$.

3. **Integral functions.** It follows from M 3 that the simple function x^n is uniquely defined for any element x of D when n is a positive integer. An integral function of x in D is then defined as the sum of a finite number of elements of D and a finite number of terms of the form $a_i x^{n_i}$ where $a_i\,(i=1,2,\ldots)$ are given elements and x is a variable element of D and $n_i\,(i=1,2,\ldots)$ are positive integers. Since $ax^m+bx^m=(a+b)x^m$, if follows that any integral function can be expressed in the form

$$a_0x^n+a_1x^{n-1}+\ldots+a_n$$

where n is a positive integer and $a_i\,(i=1,2,\ldots n)$ are elements of D. Since, however, subtraction is not necessarily always possible, we must be careful in defining exactly what is meant by saying that two integral functions are identically equal. The definition we shall use is that *two integral functions of the same variable x are said to be identically equal if, and only if, they have the same value for every value of x in D.* For instance, let us suppose that D consists only of the elements 0 and 1 whose existence is postulated in A 5 and M 5. It is readily seen that our postulates lead to $1+1=0$, that is, D is the field of all integers reduced modulo 2 which is usually denoted by $GF[2]$. In this set $x^2=x$ since $0^2=0$ and $1^2=1$. That two functions are identically equal according to our definition does not therefore necessarily imply that they differ merely in form; in other words, functions that are identically equal in one semi-field may not be so in another.

The set P of all integral functions of x in D has some, but not necessarily all, of the properties of D itself. We have in fact the following theorem if we regard functions as equal which are identically equal in the sense defined above.

THEOREM 3.1. *The set P satisfies* A 1, 2, 3, 4, 5, M 1, 2, 3, 5, AM 1 *but not necessarily* M 4 *or* AM 2.

The second part of the theorem follows from the example $GF[2]$ given above; for in this set $x \cdot x \equiv x \cdot 1$ whereas $x \not\equiv 1$. The proof of the first part is almost immediate and therefore need not be given in detail. For instance in A 4, if $a(x)$, $b(x)$ and $c(x)$ are elements of P, i. e., integral functions of x in D, and if

$$a(x) + b(x) \equiv a(x) + c(x),$$

then for every particular value of x in D the corresponding values of $a(x)$, $b(x)$ and $c(x)$ are elements of D and therefore by A 4, as applied to D, $b(x) = c(x)$ for every value of x in D; by the definition of identity we then have $b(x) \equiv c(x)$ so that these functions are equal as elements of P.

The following theorems give conditions under which two integral functions are identically equal.

THEOREM 3.2. *If two linear functions in D are identically equal, corresponding coefficients are equal.*

For if $ax + b \equiv cx + d$, then putting $x = 0$ we get $b = d$; therefore $ax \equiv cx$. Hence, putting $x = 1$, it follows that also $a = c$ as required.

THEOREM 3.3. *If $a(x)$ and $b(x)$ are two integral functions of x in D which are identically equal, then either corresponding coefficients are equal or there is an identity in D of the form $x^p \equiv x$. In the latter case there is an identity of minimum degree, $x^{m+1} \equiv x$, and, when this is used to reduce the degrees of $a(x)$ and $b(x)$ to less than $m + 1$, then corresponding coefficients of the reduced functions are equal.*

Since by hypothesis $a(x) \equiv b(x)$ then $a(0) = b(0)$; the constant terms of a and b may therefore be cancelled by A 4 as in Theorem 3.2 so that the identity may be assumed to be of the form

$$x^r(a_0 + a_1 x + \cdots + a_m x^m) \equiv x^r(b_0 + b_1 x + \cdots + b_n x^n)$$

in which a_0 and b_0 are not both zero. It follows from M 4 that

$$(3.1) \qquad a_0 + a_1 x + \cdots + a_m x^m = b_0 + b_1 x + \cdots + b_n x^n.$$

when $x \neq 0$. If $a_0 = b_0$, this equation is also true when $x = 0$ and is therefore an identity which may be treated in the same way as the original one.

If corresponding coefficients in (3.1) are not equal, there will be an equation of this form, true for all values except $x = 0$, in which the degree has a minimum value and in which $a_0 \neq b_0$; we may also take $m = n$ if we agree that either a_m or b_m may be zero but not both and we may also assume $a_m \neq 0$ without any loss of generality.

If D consists of the elements 0 and 1 only, we have seen that $x^2 = x$ and, the reduced functions being then of the first degree, corresponding coefficients are equal by Theorem 3.2. Excluding this case, let y be any element of D other than 0 or 1; we may then replace x in (3.1) by yx giving

$$a_0 + a_1 yx + \cdots + a_m y^m x^m = b_0 + b_1 yx + \cdots + b_m y^m x^m, \quad (x \neq 0;\ y \neq 0,1),$$

which, with (3.1), gives

$$\begin{aligned} &a_0 + a_1 yx + \cdots + a_m y^m x^m + y^m (b_0 + b_1 x + \cdots + b_m x^m) = \\ (3.2)\quad &b_0 + b_1 yx + \cdots + b_m y^m x^m + y^m (a_0 + a_1 x + \cdots + a_m x^m), \quad (x \neq 0;\ y \neq 0,1). \end{aligned}$$

The term of the mth degree is the same on both sides of this equation, namely $(a_m y^m + b_m y^m)x^m$, and hence this term may be cancelled. But this gives an equation of the same form as (3.1) only of lower degree, and this is only possible under our hypothesis if all corresponding coefficients in (3.1) are equal. In particular

$$a_0 + y^m b_0 = b_0 + y^m a_0.$$

Since we have assumed $a_0 \neq b_0$, it follows from AM 2 that $y^m = 1$, $(y \neq 0, 1)$. This equation is also true for $y = 1$ and therefore $y^{m+1} = y$ for all values of y, that is $x^{m+1} = x$. This completes the proof of the theorem.

It will be shown later that only semi-fields which have a finite number of elements can have an identity of the form $x^{m+1} = x$ and we shall therefore in anticipation call them finite semi-fields. In such sets the operations of addition and multiplication always have inverses except in the latter case as regards the element 0. For $x^m = 1$ may be written

$$x(x^{m-1} + x^{m-2} + \cdots + x + 1) = 1(x^{m-1} + x^{m-2} + \cdots + x + 1)$$

and therefore by M 4, if $x \neq 0, 1$,

$$x^{m-1} + x^{m-2} + \cdots + x + 1 = 0.$$

Hence if $y = x^{m-1} + \cdots + x$, then $y + 1 = 0$ and if a is any element of D and $b = ya$,

$$a + b = a + ya = (1 + y)a = 0a = 0.$$

Also if $x \neq 0$, then $x^m = 1$ so that x^{m-1} is the inverse of x with regard to multiplication. We have therefore the following theorem.

THEOREM 3.4. *If a is any element of a finite semi-field D there exists an unique element b such that $a + b = 0$, and, if $a \neq 0$, a unique element c such that $ac = 1$.*

We shall now show how to form from the original set D a new set in which subtraction and division (except by 0) is always possible. In doing so we shall assume that there is no identity $x^{m+1} \equiv x$ in the set since Theorem 3.4 shows that it is only in this case that extension may be necessary.

4. The introduction of negative elements. The theory of the division transformation* may be discussed in a perfectly general fashion in the set D but the results are somewhat cumbrous since it must not be assumed that either subtraction or division is uniformly possible in D. As for the present we shall only need to divide by the simple function $x + 1$, it is not necessary to go into the theory in detail, the following theorem being sufficient for our purposes.

THEOREM 4.1. *If $f(x)$ is any integral function in D, there exist two integral functions in D, $p(x)$ and $q(x)$, and a linear function $ax + b$ such that*

$$f(x) + p(x)(x + 1) \equiv ax + b + q(x)(x + 1).$$

For

$$\begin{aligned} x^2 + (x+1) &\equiv 1 + x(x+1), \\ x^3 + x(x+1) &\equiv x + x^2(x+1), \\ x^4 + x^2(x+1) &\equiv x^2 + x^3(x+1) \\ &\equiv 1 + (x^3 + x)(x+1), \end{aligned}$$

* By the division transformation we mean the theorem in ordinary algebra that, if $f(x)$ and $g(x)$ are polynomials of degree m and n ($m \geqq n$) there exist unique polynomials $q(x)$ and $r(x)$ such that (i) $f(x) = q(x)g(x) + r(x)$, (ii) the degree of $q(x)$ is $m - n$ and that of $r(x)$ is not greater than $n - 1$. If subtraction is sometimes inadmissable, the theorem is no longer true, e.g., if D is the semi-field of positive integers and zero, x^2 cannot be put in the form $(ax + b)(x + 1) + c$ as this requires $a = 1$, $b + a = 0$, $b + c = 0$ while there is no element in D for which $b + 1 = 0$.

and so on. It is then readily proved by induction that we have identities of the form

$$x^{2n} + p_n(x)(x+1) = 1 + q_n(x)(x+1),$$

$$x^{2n+1} + xp_n(x)(x+1) = x + xq_n(x)(x+1)$$

for all integral values of n, from which the theorem is immediately apparent. We shall say that $f(x)$ is *congruent to* $ax+b$ *modulo* $(x+1)$ or

$$f(x) \equiv ax + b \pmod{x+1},$$

and also that $ax+b$ *is equal to* $f(x)$ *reduced modulo* $(x+1)$. Thus

$$x^n \equiv 1 \text{ mod. } (x+1) \qquad (n \text{ even}),$$

$$\equiv x \text{ mod. } (x+1) \qquad (n \text{ odd}).$$

Corollary. If the integral functions $f(x)$ *and* $g(x)$ *are congruent to the same linear function modulo* $x+1$, *there exist integral functions* $P(x)$ *and* $Q(x)$ *such that*

$$f(x) + P(x)(x+1) = g(x) + Q(x)(x+1).$$

For, if

$$f(x) + p_1(x)(x+1) = ax + b + q_1(x)(x+1),$$

$$g(x) + p_2(x)(x+1) = ax + b + q_2(x)(x+1),$$

then

$$f(x) + (p_1(x) + q_2(x))(x+1) = ax + b + (q_1(x) + q_2(x))(x+1)$$

$$= g(x) + (p_2(x) + q_1(x))(x+1).$$

Integral functions which are congruent to the same linear function are said to be congruent to each other modulo $x+1$. If $f(x) \equiv g(x)$ and $g(x) \equiv h(x)$ mod. $(x+1)$, then evidently also $f(x) \equiv h(x)$ mod. $(x+1)$. All integral functions in D which are congruent to the same function $f(x)$ modulo $x+1$ form a set which is called the *class* $[f(x)]$ corresponding to $f(x)$; by Theorem 4.1

every class contains a function of the form $ax+b$ which may be taken as typical of that class as every function of the class is congruent to it. It follows immediately that the classes $[f(x)]$ and $[g(x)]$ are equal (that is, are composed of the same set of integral functions) if, and only if, $f(x) \equiv g(x)$ mod. $(x+1)$.

The classes [0] ane [1] are called the zero and identity classes. The sum and product of two classes are defined by

$$[f(x)]+[g(x)] = [f(x)+g(x)], \qquad [f(x)]\,[g(x)] = [f(x)g(x)].$$

We then have the following theorem.

THEOREM 4.2. *The set K of all classes of integral functions of x in D forms a semi-field.*

This follows immediately for A 1, 2, 3, 4, 5. M 1, 2, 3, 5. We shall therefore only give the proof of A 4 as an example of the method of demonstration.

Let α, β and γ in K correspond to $f(x)$, $g(x)$ and $h(x)$ in P; then, if $\alpha+\beta=\alpha+\gamma$, it follows from the definition of K that

$$f(x)+g(x) \equiv f(x)+h(x) \text{ mod. } (x+1).$$

There exist therefore functions $\varphi(x)$ and $\psi(x)$ in P such that

$$f(x)+g(x)+\varphi(x)\,(x+1) = f(x)+h(x)+\psi(x)\,(x+1);$$

hence

$$g(x)+\varphi(x)\,(x+1) = h(x)+\psi(x+1),$$

or

$$g(x) \equiv h(x) \text{ mod. } (x+1),$$

that is $\beta=\gamma$ as required by A 4.

Before proving M 4 and A M 2 we require the following lemmas.

LEMMA 4.1. *If a and b are elements of D and $a \equiv b$ mod. $(x+1)$, then $a=b$.*

From the definition of congruence, we have

$$a+p(x)\,(x+1) = b+q(x)\,(x+1). \tag{4.1}$$

If D is finite, there is by Theorem 3.4 a value x_1 of x for which $x_1+1=0$ and inserting this value in (4.1) we get $a=b$.

If D is not finite, the coefficients of corresponding powers of x in (4.1) are equal. Writing (4.1) in the form

$$a+(p_0+p_1 x+\cdots p_m x^m)(x+1) = b+(q_0+q_1 x+\cdots q_m x^m)(x+1)$$

we have on expansion

$$p_m x^{m+1}+(p_m+p_{m-1})x^m+\cdots+(p_1+p_0)x+p_0+a$$
$$= q_m x^{m+1}+(q_m+q_{m-1})x^m+\cdots+(q_1+q_0)x+q_0+b,$$

and on equating coefficients we readily derive $a = b$.

LEMMA 4.2. *If* $ax+b \equiv cx+d$ mod. $(x+1)$, *then* $a+d = b+c$ *and conversely.*

For if $ax+b \equiv cx+d$ mod. $(x+1)$, then

$$ax+b+c(x+1) \equiv cx+d+a(x+1) \text{ mod. } (x+1),$$

or

$$(a+c)x+b+c \equiv (a+c)x+a+d \text{ mod. } (x+1),$$

and hence, as in the proofs of A 4 and Lemma 4.1 above, $b+c = a+d$.

Conversely, if $b+c = a+d$, then

$$(a+c)x+b+c \equiv (a+c)x+a+d,$$

or

$$ax+b+c(x+1) \equiv cx+d+a(x+1),$$

that is

$$ax+b \equiv cx+d \text{ mod. } (x+1).$$

Returning now to AM 2, let $\alpha, \beta, \gamma, \delta$ be elements of K such that

$$\alpha\gamma+\beta\delta = \alpha\delta+\beta\gamma, \qquad \alpha \neq \beta, \tag{4.2}$$

and $\alpha_0 x + \alpha_1$, $\beta_0 x + \beta_1$, $\gamma_0 x + \gamma_1$, $\delta_0 x + \delta_1$ corresponding elements of P so that

$$(4.3)\qquad \begin{aligned}&(\alpha_0 x + \alpha_1)(\gamma_0 x + \gamma_1) + (\beta_0 x + \beta_1)(\delta_0 x + \delta_1)\\ &\qquad \equiv (\alpha_0 x + \alpha_1)(\delta_0 x + \delta_1) + (\beta_0 x + \beta_1)(\gamma_0 x + \gamma_1) \bmod.(x+1)\end{aligned}$$

and, since $\alpha \neq \beta$, it follows from Lemma 4.2 that

$$(4.4)\qquad \alpha_0 + \beta_1 \neq \alpha_1 + \beta_0.$$

Expanding (4.3) and using $x^2 \equiv 1 \bmod.(x+1)$, we get

$$\begin{aligned}&(\alpha_0\gamma_1 + \alpha_1\gamma_0 + \beta_0\delta_1 + \beta_1\delta_0)x + \alpha_0\gamma_0 + \alpha_1\gamma_1 + \beta_0\delta_0 + \beta_1\delta_1\\ &\quad \equiv (\alpha_0\delta_1 + \alpha_1\delta_0 + \beta_0\gamma_1 + \beta_1\gamma_0)x + \alpha_0\delta_0 + \alpha_1\delta_1 + \beta_0\gamma_0 + \beta_1\gamma_1 \bmod.(x+1)\end{aligned}$$

and therefore by Lemma 4.2

$$\begin{aligned}&\alpha_0\gamma_1 + \alpha_1\gamma_0 + \beta_0\delta_1 + \beta_1\delta_0 + \alpha_0\delta_0 + \alpha_1\delta_1 + \beta_0\gamma_0 + \beta_1\gamma_1\\ &\qquad = \alpha_0\delta_1 + \alpha_1\delta_0 + \beta_0\gamma_1 + \beta_1\gamma_0 + \alpha_0\gamma_0 + \alpha_1\gamma_1 + \beta_0\delta_0 + \beta_1\delta_1\end{aligned}$$

or

$$(\alpha_0 + \beta_1)(\gamma_1 + \delta_0) + (\alpha_1 + \beta_0)(\gamma_0 + \delta_1) = (\alpha_0 + \beta_1)(\gamma_0 + \delta_1) + (\alpha_1 + \beta_0)(\gamma_1 + \delta_0).$$

Hence, since AM 2 holds in D and $\alpha_0 + \beta_1 \neq \alpha_1 + \beta_0$ by (4.4), we have

$$\gamma_1 + \delta_0 = \gamma_0 + \delta_1,$$

and therefore by Lemma 4.2

$$\gamma_0 x + \gamma_1 \equiv \delta_0 x + \delta_1 \bmod.(x+1),$$

that is, $\gamma = \delta$ in K; hence AM 2, and therefore $M4$ also, is true in K.

We can now prove that the operation of addition always has an inverse in K; this theorem, which is one of the defining properties of a field, we shall refer to as A 6.

A 6. *If α is any element of K, there exists a unique element β of K such that* $\alpha + \beta = 0$.

Let α in K correspond to $\alpha_0 x + \alpha_1$ in P, and let β be the element of K which corresponds to $\alpha_1 x + \alpha_0$: then

$$\alpha_0 x + \alpha_1 + \alpha_1 x + \alpha_0 = (\alpha_1 + \alpha_0)(x + 1) = 0 \text{ mod.}(x + 1).$$

Hence

$$\alpha + \beta = 0.$$

The uniqueness of β follows immediately from A 4.

5. **The inverse of multiplication; rational numbers.** We shall now derive from K a set in which division by any element except 0 is admissible using methods similar to those already used in introducing subtraction. As we have already seen that this is the case when there is in K an identity of the form $x^{m+1} = x$, we shall assume in this section that there is no such identity.

In place of the single variable x of the last paragraph we now introduce a set of variables, one to each element of K except 0, denoting the variables which correspond to the elements a, b, etc. by x_a, x_b, etc. respectively. We then consider the set KI of all integral functions of these variables with coefficients in K which contain a finite number of the variables x_a and are of finite degree. Then from KI we derive in turn a new set KR of all elements of KI reduced modulis $ax_a - 1$, $bx_b - 1$, ..., that is, two elements f and g of KI correspond to the same element of KR if, and only if, there exist a finite number of elements q_a, q_b, ... q_k of KI such that

$$f = g + q_a(ax_a - 1) + q_b(bx_b - 1) + \cdots + q_k(kx_k - 1),$$

a relation which we shall denote by writing

$$f = g \text{ modd. } (L).$$

For the sake of brevity we shall also write

$$\alpha = f \text{ modd. } (L), \text{ or } \alpha = [f]$$

if α is the element of KR corresponding to the element f of KI. Addition and multiplication are defined in KR exactly as in K. If α and β are two elements of KR and f and g corresponding elements in KI, then $\alpha+\beta$ is the element of KR which corresponds to $f+g$ in KI, and similarly $\alpha\beta$ corresponds to fg. The method of finding the element of KR which corresponds to a given element of KI is not quite so simple as in the similar problem in the previous section but depends on the following theorem due to Kronecker.

THEOREM 5.1. *Every element of KI is congruent modd. (L) to a linear function of the form αx_a.*

This follows immediately from Kronecker's identities

$$(5.1)\quad \alpha x_a+\beta x_b \equiv (\alpha b+\beta a)x_{ab}+\alpha b x_{ab}(a x_a-1)+\beta a x_{ab}(b x_b-1) - (\alpha x_a+\beta x_b)(a b x_{ab}-1)$$
$$\equiv (\alpha b+\beta a)x_{ab} \text{ modd. } (L).$$

$$(5.2)\quad x_a\, x_b \equiv x_{ab}+b x_b\, x_{ab}(a x_a-1)+x_{ab}(b x_b-1)-x_a\, x_b(a b x_{ab}-1)$$
$$\equiv x_{ab} \text{ modd. } (L).$$

and

$$\beta = \beta x_i - \beta(x_i-1) \equiv \beta x_i \text{ modd. } (L).$$

where x_i is the variable corresponding to the element 1 of KI.

It is easily seen as in § 4 that KR satisfies A 1, 2, 3, 4, 5, 6, M 1, 2, 3, 5, AM 1. To show that it satisfies M 4 we require the following lemmas.

LEMMA 5.1. *If an element of KI is identically zero, all its coefficients are zero.*

The proof is made exactly as in ordinary algebra*, and we shall therefore give it in outline only. Denoting variables in K by $x_1, x_2 \cdots$ instead of by $x_a, x_b, \cdots$ as formerly, let $\varphi(x_1, x_2, \ldots x_n)$ be any integral function in K i.e., any element of KI, which is identically zero. If there is only one variable, the lemma is the same as Theorem 3.3; we therefore assume that it is true for $n-1$ variables. Arranging φ according to powers of x_n we may write

$$\varphi = \varphi_0\, x_n^m + \varphi_1\, x_n^{m-1} + \cdots + \varphi_m,$$

* Cf. Böcher, Introduction to Higher Algebra, p. 5.

where φ_i $(i = 0, 1, \ldots, m)$ are integral functions of $x_1, x_2, \ldots, x_{n-1}$. If $x_1, \ldots x_{n-1}$ are then given fixed but arbitrary values, φ becomes a function of x_n alone and, being identically zero, its coefficients φ_i are zero. Hence $\varphi_0, \ldots \varphi_m$ vanish for all values of $x_1, \ldots x_{n-1}$. They are therefore identically zero and hence by hypothesis all their coefficients are zero. The lemma is therefore true for n variables and by induction for any number of variables.

LEMMA 5.2. *If an integral function $\varphi(x_1, x_2, \ldots x_n)$ in KI is identically zero so is also the integral function*

$$\psi(x_1, x_2, \ldots, x_n) = x_1^{m_1} x_2^{m_2} \cdots x_n^{m_n} \varphi\left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots \frac{1}{x_n}\right),$$

where m_i $(i = 1, 2, \ldots n)$ is the degree of φ in x_i.

For if

$$\varphi = \varphi_0 x_1^{m_1} + \varphi_1 x_2^{m_1 - 1} + \cdots + \varphi_{m_1} = 0,$$

then by Lemma 5.1 $\varphi_i = 0$ $(i = 0, 1, \ldots m_1)$, and hence also

$$\varphi_0 + \varphi_1 x_1 + \cdots \varphi_{m_1} x_1^{m_1} = 0.$$

The lemma therefore follows easily by induction.

LEMMA 5.3. *If $\alpha x_a = 0$ modd. (L), then $\alpha = 0$.*

If $b_1, b_2, \ldots$ are elements of K no one of which is zero, and $x_1, x_2 \ldots$ the corresponding variables in KI, there must be an identity of the form

$$\alpha x_a + \sum_j \psi_j(x_1, x_2, \ldots)(b_j x_j - 1) = 0,$$

and hence by Lemma 5.2

$$\alpha x_a^m x_1^{m_1} x_2^{m_2} \cdots + \sum_j x_a^{m^{(j)}} x_1^{m_1^{(j)}} \cdots \psi_j\left(\frac{1}{x_1}, \frac{1}{x_2}, \ldots\right)(b_j - x_j) = 0$$

where the exponents m depend on the degrees of $\psi_1, \psi_2, \ldots$ in the respective variables. Putting $x_a = a$, $x_j = b_j$ $(j = 1, 2, \ldots)$ in this identity we get

$$\alpha a^m b_1^{m_1} b_2^{m_2} \cdots = 0$$

and therefore $\alpha = 0$.

Suppose now that a, b, c $(a \neq 0)$ are elements of KR, that is, integral functions reduced modd (L) by means of Theorem 5.1, and suppose that $ab = ac$ or, which the same thing, $a(b-c) = 0$. Let the elements of KI which correspond to a and $b-c$ be ax_p $(\alpha \neq 0)$ and βx_q respectively; then

$$0 \equiv \alpha x_p \, \beta x_q \equiv \alpha\beta x_{pq} \text{ modd. } (L)$$

so the $\alpha\beta = 0$ by Lemma 5.3. Since $\alpha \neq 0$, we have $\beta = 0$ so that $b - c = 0$ or $b = c$; M 4 is therefore true in KR.

We can now prove the following theorem which we shall refer as M 6.

M 6 *Every element of KR, except 0, has an inverse.*

For

$$ax_b \cdot bx_a \equiv 1 + (bx_b - 1) + bx_b(ax_a - 1)$$
$$\equiv 1 \text{ modd. } (L).$$

Collecting the results of the preceding paragraphs we have

THEOREM 5.2. *The set KR satisfies* A 1, 2, 3, 4, 5, 6, M 1, 2, 3, 4, 5, 6, AM 1.

Any set for which this theorem is true is called a *field* or domain of rationality. A field then possesses all the properties that are necessary for the validity of any algebraic theorem* which does not depend on the idea of order or as we shall say, any theorem of rational algebra; and this is true whether the field is finite or not, if care is taken in the finite case to reduce the degree of every integral function to less than $m+1$ by means of the identity $x^{m+1} \equiv x$.

We shall state here for reference the most important of these theorems but shall leave the proofs to the reader. For the sake of brevity we shall assume that all the functions used are integral functions whose coefficients are elements of a field F that is to say integral functions rational in F.

THEOREM 5.3. *If the degree of $f(x)$ is n, f cannot vanish for more than n different values of x.*

Here the degree must be reduced to m or less if F is finite.

Definition. An integral function rational in F is said to be reducible if it is the product of two integral functions in F neither of which is a constant.

THEOREM 5.4. *$f(x)$ is expressible in essentially only one way as the product of factors which are rational and irreducible in F.*

THEOREM 5.5. *If $f(x)$ and $g(x)$ are of degree m and n respectively $(m \geqq n)$ there exist unique functions $q(x)$ and $r(x)$, in which the degree of r is less than n, such that $f(x) \equiv q(x)\, g(x) + r(x)$.*

* In modular fields (§ 6) care must be taken that division by 0 does not occur; cf. second footnote on p. 261.

THEOREM 5.6. *If $h(x)$ is the highest common factor of $f(x)$ and $g(x)$, $h(x)$ is rational in F and there exist in F integral functions $\varphi(x)$ and $\psi(x)$, whose degrees are not greater than $n-1$ and $m-1$ respectively, such that $\varphi(x)f(x) + \psi(x)g(x) \equiv h(x)$.*

THEOREM 5.6. *If $f(x)$ is irreducible in F and $g(x)$ is a rational integral function which has a factor in common with $f(x)$, then $g(x)$ is exactly divisible by $f(x)$.*

6. **Modular fields and finite fields.** A field F is said to be modular if for some element $a \neq 0$ there are two integers m and n such that* $ma = na$ or $(m-n)a = 0$. Let $pa = 0$ be the identity of this form for which p has its smallest value. Since $pa = (p \times 1)a$, it follows from M4 that $p \times 1 = 0$. Hence, if x is any element of F,

$$px = (p \times 1)x = 0;$$

the value of p is therefore the same for all elements of F except 0. It is easily seen that p is a prime; for if $p = mn$, then $(m \times 1)(n \times 1) = 0$ and therefore either $m \times 1 = 0$ or $n \times 1 = 0$, and, as p is minimal, either m or n must equal p. Further any integer q such that $qa = 0$ is divisible by p; for if this is not the case there exist integers l and m such that $lp + mq = 1$ and therefore

$$a = (lp + mq)a = lpa + mqa = 0.$$

These results may be summarized as follows.

THEOREM 6.1. *If one element of a field is modular, all elements are modular and all have the same modulus; further the modulus is a prime.*

As an elementary example of a modular field we may take the set of integers reduced modulo 3. In this field there are three elements 0, 1, and 2 such that

$$1+1 = 2,\ 2+2 = 1,\ 1+2 = 0,\ 1+1+1 = 0,\ 2+2+2 = 0,\ 2^2 = 1.$$

Every element satisfies the identity $x^3 = x$. Another example is the field with four elements $0, 1, a, a^2$ in which

$$1+1 = a+a = a^2+a^2 = 0, \qquad a^3 = 1, \qquad a^2 = a+1.$$

In this field we have the identity $x^4 = x$.

* Here ma is used as a contraction for the result of adding a to itself $m-1$ times, e.g. $a + a = 2a$. It is not implied that m is an element of the field, although we shall on occasion use m to denote $m \times 1$. It is evident that $ma = (m \times 1)a$.

It is clear that every field with only a finite number of elements must be modular; for if this were not so we could generate an infinity of elements by the repeated addition of 1 to itself. Further, if any element $a \neq 0$ is repeatedly multiplied into itself, we must at some stage arrive at the value 1; for as there are only a finite number of elements there must exist integers α and β for which $a^\alpha = a^\beta$ and therefore $a^{\alpha-\beta} = 1$. If m is the least common multiple of the exponents $\alpha - \beta$ for all non-zero elements of F, then for any such element $x^m = 1$, and therefore for every such element, including 0, we have $x^{m+1} = x$. Conversely, if $x^{m+1} = x$ in F, the number of elements in F is necessarily finite; for every element except 0 is a solution of the equation $x^m = 1$ and this equation cannot have more than m roots by Theorem 5.3. Further $m+1$ is the number of elements in F. For if the elements of F are $0, a_i$ $(i = 1, 2, \ldots)$ then the integral function

$$\varphi(x) = x \prod_i (x - a_i)$$

vanishes identically, or

$$\psi(x) = \prod_i (x - a_i) = 0$$

for every value of x except $x = 0$. Hence, as m is minimal, the degree of $\varphi(x)$ must be $m+1$ and in fact it is easily shown that $\psi(x) = x^m - 1$. We have therefore the following theorem.

THEOREM 6.2. *Any field which is finite is modular and, if μ is the number of elements in the field, $x^\mu = x$ for every value of x; conversely every field in which there is an identity of this form is finite.**

The lowest power ν of a given element x of F which equals 1 is called the *period* of x. Let a be an element for which the period ν has its maximum value. If $x^\nu \neq 1$ for some value of $x \neq 0$, there must be some element b whose period σ is less than ν but not a factor of it; also, if ϱ is the G. C. M. of ν and σ and $\sigma = k\varrho$, the period of b^ϱ is k and b^ϱ is therefore an element whose period is relatively prime to ν. We shall therefore assume that σ is relatively prime to ν. If τ is the period of ab, then

$$1 = (ab)^{\tau\sigma} = a^{\tau\sigma},$$

* E. H. Moore, A doubly infinite system of simple groups. Math. Papers, Internat. Math. Congress (1893), p. 211.

so that $\tau\sigma$, and therefore also τ, is divisible by ν; but

$$(ab)^{\nu\sigma} = a^{\nu\sigma} b^{\nu\sigma} = 1;$$

hence the period of ab is $\nu\sigma$ and $\tau = \nu$. This period is however greater than ν, contrary to our hypothesis, and hence the period of every element (except 0) is a factor of ν. This means that there are m different powers of a and we therefore have the following theorem.

THEOREM 6.3. *If F is a finite field, there is an element a of F whose period is $\mu - 1$ and all elements of F except 0 are integral powers of a.*

This means of course that the non-zero elements form a cyclic group of order $\mu - 1$.

We shall now prove the following theorem.

THEOREM 6.4. *The number μ of elements in a finite field F is a power of the modulus p.*

The field F, if it exists, must contain elements congruent to $0, 1, 2, \ldots p-1$ mod. p and it is in fact easily verified that these elements themselves form a field. Let P denote this subset of F and p_i $(i = 1, 2, \ldots p)$ its elements. If x_1 is an element of F which is not contained in P, the p^2 elements $p_i + p_j x_1$ $(i, j = 1, 2, \ldots p)$ are all different; for were $p_i + p_j x_1 = p_k + p_l x_1$, then $(p_j - p_l) x_1 = p_k - p_i$ and hence either $p_j = p_l$ and therefore $p_k = p_i$ or $x_1 = (p_j - p_l)^{-1}(p_k - p_i)$, which is impossible since x_1 is not contained in P. In the same way, if x_2 is not contained in the set P_2 of all elements of the form $p_i + p_j x_1$, the set P_3 of elements of the form

$$p_i + p_j x_1 + p_k x_2 \qquad (i, j, k = 1, 2, \ldots p)$$

are p^3 in number and all different; and so on. As there are only a finite number of elements in F, this process must eventually come to an end after say n steps, and there exist therefore n elements, $1, x_1, x_2, \ldots x_{n-1}$, such that every element can be uniquely expressed in the form

$$x = p_{i_0} + p_{i_1} x_1 + \cdots + p_{i_{n-1}} x_{n-1} \qquad (i_0, \ldots i_{n-1} = 1, 2, \ldots p),$$

so giving p^n different elements.*

The elements $1, x_1, x_2, \ldots x_{n-1}$ are said to form a *basis of F relatively to P*.

* For a fuller discussion of finite fields see L. E. Dickson, Linear Groups, Leipzig 1901.

7. The extension of a field by adjunction. As was pointed out in the introduction, the idea on which the various extensions of the original set D is based is that, whenever an algebraic problem in a given set leads to an equation which has no solution, a new set is defined by algebraic processes in which this equation does have a solution. In §§ 4 and 5 the exposition of the method was somewhat cramped owing to the possibility that in some cases subtraction and division of polynomials might be impossible; it is therefore necessary now to consider the details of the method as applied to a field somewhat more closely.

We suppose a field F given as a starting point; this field may by infinite or finite. We then consider the set F_x of all rational integral functions of x in F, that is, all integral functions of x of finite degree whose coefficients are elements of F.

If $\varphi(x)$ is an element of F_x of degree n, and $f(x)$, $g(x)$ and $h(x)$ are also elements of this set such that

$$f(x) = g(x) + h(x)\,\varphi(x) \tag{7.1}$$

then $f(x)$ is said to be congruent to $g(x)$ modulo $\varphi(x)$ or

$$f(x) \equiv g(x) \text{ mod. } \varphi(x).$$

Since the division transformation (Theorem 5.5) is always possible in a field, every element of F_x is congruent to a unique element $g(x)$ if the restriction is added in (7.1) that the degree of $g(x)$ is to be $n-1$ at most.

We then form a new set, which we shall denote by $F[\varphi(x)]$, the elements of which are in (1, 1) correspondence with those elements of F_x whose degree does not exceed $n-1$, and in this set we define the operations of addition as follows. If f and g are elements of $F[\varphi(x)]$ and $f(x)$ and $g(x)$ the corresponding elements of F_x of degree $n-1$ or less, then

(i) $f+g$ is defined as the element of $F[\varphi(x)]$ which corresponds to $f(x)+g(x)$ in F_x;

(ii) fg is defined as the element of $F[\varphi(x)]$ which corresponds to the remainder in F_x when $f(x)\,g(x)$ is divided by $\varphi(x)$.

We may without chance of confusion write $f \equiv f(x)$ mod. $\varphi(x)$, $f+g \equiv f(x)+g(x)$ mod. $\varphi(x)$, and $fg \equiv f(x)\,g(x)$ mod. $\varphi(x)$ or in other words, the set $F[\varphi(x)]$ is what the set F_x becomes when polynomials which are congruent modulo $\varphi(x)$ are regarded as equivalent; for this reason we shall often refer to the set $F[\varphi(x)]$ as the set F_x reduced modulo $\varphi(x)$.

The elements of $F[\varphi(x)]$ which correspond to elements of F_x of degree 0, that is, to elements of F itself, form a field simply isomorphic with F; we shall as a rule not distinguish between these fields and will refer to both as the field F. This will be found to lead to no confusion. We shall now prove the following fundamental theorem.

THEOREM 7.1. *If $\varphi(z)$ is an irreducible* integral function of z in a field F, the set $F[\varphi(z)]$ of integral functions of z reduced modulo $\varphi(z)$ forms a field.*

The only one of the postulates for a field which is not obviously satisfied is M4 and to prove that this postulate holds it is only necessary to show that

$$f(z)\,g(z) \equiv 0 \text{ mod. } \varphi(z) \tag{7.2}$$

entails either $f(z) \equiv 0$ or $g(z) \equiv 0$ mod. $\varphi(z)$.

Suppose $f(z) \not\equiv 0$ mod. $\varphi(z)$. Since $\varphi(z)$ is irreducible, f and φ have no common factor and there are therefore polynomials $h(z)$ and $k(z)$ in F_z such that

$$h(z)f(z) + k(z)\varphi(z) = 1,$$

that is

$$h(z)f(z) \equiv 1 \text{ mod. } \varphi(z),$$

or $h(z)$, is the inverse of $f(z)$ modulo $\varphi(z)$. It follows that $g(z) \equiv h(z)f(z)g(z)$ mod. $\varphi(z)$ and therefore from (7.2) $g(z) \equiv 0$ mod. $\varphi(z)$, as required by the postulate.

If α is the element of $F[\varphi(z)]$ which corresponds to the integral function z in F_z, then, since $\varphi(z) \equiv 0$ mod. $\varphi(z)$, we have $\varphi(\alpha) = 0$. Hence $x - \alpha$ is a factor of $\varphi(x)$ in $F[\varphi(z)]$ and so $\varphi(x)$ is reducible in the new field.

Every element of $F[\varphi(z)]$ evidently has the form

$$x = a_0 + a_1\alpha + a_2\alpha^2 + \ldots a_{n-1}\alpha^{n-1}$$

where a_i $(i = 0, 1, \ldots n-1)$ are elements of F (regarded as a subset of the new field). The field $F[\varphi(z)]$ is generally denoted by $F(\alpha)$ and it is said to be derived from F by the *adjunction* of α to F. $F(\alpha)$ is said to be an algebraic field *over* F, and F a *subfield* of $F(\alpha)$. The meaning of $F(\alpha)$ is of course indefinite till the corresponding irreducible function $\varphi(z)$ is given.

* We shall tacitly assume that only irreducible functions of higher degree than the first are used as, when this is not the case, the new set is the same as the old.

If the degree of $\varphi(z)$ is n, the field $F(\alpha)$ is also said to be of *degree n with respect to F.*

8. **The field as a linear associative algebra.** We saw in the previous section that any element of $F[\varphi(z)]$ is linearly expressible in terms of $1, \alpha, \alpha^2, \ldots, \alpha^{n-1}$ with coefficients in F. In place of the powers of α we may take any n linearly independent elements of $F(\alpha)$, say $e_1, e_2, \ldots, e_n$, that is, any n elements such that there is no relation of the form $\sum_i n_i e_i = 0$ where the n's are elements of F which are not all zero. If the e's are so chosen, it follows from the theory of linear independence that any element x of $F(\alpha)$ can be expressed uniquely in the form

$$(8.1) \qquad x = \xi_1 e_1 + \xi_2 e_2 + \cdots + \xi_n e_n$$

where the ξ's are elements of F and in particular the product of any two e's has this form, say

$$(8.2) \qquad e_i e_j = \sum_{k=1}^{n} \gamma_{ijk} e_k,$$

where, since $e_i e_j \cdot e_k = e_i \cdot e_j e_k$,

$$(8.3) \qquad \sum_\lambda \gamma_{ij\lambda} \gamma_{\lambda k\mu} = \sum_\lambda \gamma_{i\lambda\mu} \alpha_{jk\lambda};$$

and also, since $e_i e_j = e_j e_i$,

$$(8.4) \qquad \gamma_{ijk} = \gamma_{jik}.$$

These equations (8.1, 2, 3, 4) may be taken as defining the field together with a further restriction on the γ's to ensure that every element except 0 has an inverse.

Kronecker's method of regarding this definition is as follows. He sets $f_{ij} = e_i e_j - \sum_k \gamma_{ijk} e_k = f_{ji}$, the e's being variables in F, and reduces the set of all rational integral functions of the e's modulis f_{ij} $(i, j = 1, 2, \ldots n;\ i \leq j)$. The reader may easily show that this leads to a field if conditions are imposed on the γ's so as to render multiplication associative and invertible. Hamilton on the other hand regards the field as the set of all sets $(\xi_1, \xi_2, \ldots \xi_n)$ of n ordered elements of F. Two such sets $x = (\xi_1, \xi_2, \ldots \xi_n)$ and $y = (\eta_1, \eta_2, \ldots \eta_n)$ are said to be equal if, and only if, $\xi_i = \eta_i$ $(i = 1, 2, \ldots n)$; the sum $x + y$

of the two sets is defined as the set $(\zeta_1, \zeta_2, \ldots \zeta_n)$ where $\zeta_i = \xi_i + \eta_i$, and the product is defined as $xy = (\lambda_1, \lambda_2, \ldots \lambda_n)$ where $\lambda_k = \sum_{ij} \gamma_{ijk} \xi_i \eta_j$.

Any set of elements which satisfies the postulates for addition and also (8.2, 3) is called a *linear associative algebra* which is said to be commutative if (8.4) is also true. Any set of linearly independent elements, such as $e_1, e_2, \ldots e_n$, is said to form a *basis* of the algebra, and the number of elements in a basis is called the order* of the algebra.

If $\psi(\alpha, z)$ is an integral function of z which is rational and irreducible in $F(\alpha)$, we can extend $F(\alpha)$ by taking the set of all integral functions of z in $F(\alpha)$ reduced modulo $\psi(\alpha, z)$, just as previously we extended F to form $F(\alpha)$. It follows then that, if β corresponds to z as a polynomial in $F(\alpha)$, every element of the extended field has the form $\sum_{ij} \xi_{ij} \alpha^i \beta^j$, and the field is again a linear associative algebra of order† mn.

Any field derived from F by a finite number of such steps is called an *algebraic field over* F. Every algebraic field over F is therefore a linear associative algebra in F.

9. **The identical equation.** Let F' be an algebraic field of degree n over F, and let a basis of it be $e_1, e_2, \ldots e_n$. A variable element x of F' may be written

$$x = \xi_1 e_1 + \xi_2 e_2 + \cdots + \xi_n e_n$$

where the ξ's are variable elements of F. If we form the first $n + 1$ powers of x, namely $x^0 = 1, x, x^2, \ldots x^n$, we have $n + 1$ elements of F' which cannot be linearly independent as the greatest number of linearly independent elements in F' is n. There must therefore exist a relation of the form

$$g(x) = b_0 x^n + b_1 x^{n-1} + \cdots + b_{n-1} x + b_n = 0, \tag{9.1}$$

where the b's are rational functions of the coefficients ξ which may be taken to be integral, and the equation is an identity in the ξ's when x^i $(i = 1, 2, \ldots n)$ are expressed linearly in terms of $e_1, e_2, \ldots e_n$.

The coefficients b_i may be taken to be homogeneous functions of the ξ's of degree $\mu - n + i$ where μ is the degree of g as a function of the ξ's; for if we replace x by λx in (9.1), it becomes an identity in $\lambda, \xi_1, \ldots \xi_n$, and, if

$$b_i = \lambda^{r_i} b_{i0} + \lambda^{r_i - 1} b_{i1} + \cdots, \qquad (r_i = \mu - n + i),$$

* In the type of algebra $F(\alpha)$ considered above the "order" is the same as the "degree".

† If the order were less than mn, β would satisfy an equation in $F(\alpha)$ of lower degree than m which is impossible since $\psi(\alpha, z)$ is irreducible in $F(\alpha)$.

then the coefficient of the highest power of λ must vanish identically in the ξ's, and this coefficient is the homogeneous expression $\sum_i b_{i^0} x^{n-i}$. It is conceivable of course that x might satisfy an equation of lower degree than n. Assume that the equation of lowest degree is

$$(9.2) \qquad f(x, \xi) = a_0(\xi) x^m + a_1(\xi) x^{m-1} + \cdots + a_m(\xi) = 0,$$

and let f be of the νth degree* in the ξ's (which occur both explicitly and in the powers of x) so that a_i is a homogeneous polynomial in the ξ's of degree $\nu - m + i$. This equation we shall call the *identical equation* of the field F' in F.

The condition that $1, x, \ldots x^{m-1}$ are linearly independent while $1, x, \ldots x^{m-1}$, x^m are linearly dependent can be expressed as the non-vanishing or vanishing, respectively, of certain determinants, hence there are always particular values of the ξ's for which x satisfies no equation in F of lower degree than m. For such an element $f(t, \xi)$ is an irreducible function of t in F, from which it follows in particular that f has no repeated roots and therefore† its derivative $f'(t, \xi)$ with regard to t does not vanish when x is substituted for t. An element x which satisfies no equation of lower degree than m is called a *primitive* element of F'.

Let x be a primitive element, $y = \sum_i \eta_i e_i$ a variable element of F', and λ a variable in F; then

$$(9.3) \qquad f(x + \lambda y, \xi + \lambda \eta) = 0$$

is an identity in $\lambda, \xi_1, \ldots \xi_n, \eta_1, \ldots \eta_n$. If (9.3) is expanded in powers of λ, the coefficient of each power must be identically zero in the ξ's and η's. Hence, if we set

$$a_i(\xi + \lambda \eta) = a_i(\xi) + \lambda a_{i1}(\xi, \eta) + \cdots$$

and equate the coefficient of the first power of λ to 0, we have

$$f'(x, \xi) y + \sum_i a_{i1}(\xi, \eta) x^{m-i} = 0.$$

* It can be shown that $a_0 = 1$ and therefore $\nu = m$, but we shall have no need of this fact; f is written $f(x, \xi)$ to emphasize the fact that ξ occurs in the coefficients as well as in the powers of x.

† This is not necessarily the case in modular fields; cf. Steinitz, loc. cit., p. 212.

Since x is primitive, $f'(x, \xi) \neq 0$, and it follows that

$$y = -\frac{a_{i1}(\xi, \eta)\, x^{m-1}}{f'(x, \xi)} = \chi(x).$$

We may take χ as an integral function of x. For, if $g(x)$ is any polynomial in x which is not zero, $g(t)$ and $f(t, \xi)$ have no common factor and hence we can find two rational polynomials $p(t)$ and $q(t)$ such that $p(t)g(t) + q(t) f(t, \xi) = 1$; hence $p(x)\, g(x) = 1$ so that $1/g(x) = p(x)$, a polynomial in x.

Since y is arbitrary, it follows that every element of F' is linearly dependent on $1, x, \ldots x^{m-1}$ so that m is the degree of F'. We have therefore the following theorem.

THEOREM 9.1. *Every element of F' satisfies an equation $f(x, \xi) = 0$ in F of degree n and there are elements, called primitive elements, which satisfy no rational equation in F of lower degree. Further, if x is a primitive element, every element of F' can be expressed as rational polynomial of x in F, that is, $F' = F(x)$.*

10. **Subfields.** Let $F(\alpha)$ be an algebraic field of degree n over F and F' a subfield of $F(\alpha)$ which contains F.

THEOREM 10.1. *Any subfield of $F(\alpha)$ which contains F is an algebraic field over F.*

Let us denote elements of F by the letter ξ with appropriate affixes. If ϵ_1 is any element of F' which is not in F, then 1 and ϵ_1 are linearly independent in F; for, were $\xi_0 + \xi_1 \epsilon_1 = 0\ (\xi_1 \neq 0)$, then $\epsilon_1 = -\xi_0/\xi_1 < F$. We shall denote the set of all elements of F' which are linearly dependent on 1 and ϵ_1 in F by $A_1 = (1, \epsilon_1)$; or more generally if $x_1, x_2, \ldots x_r$ are any r elements of F', the set of elements linearly dependent on these elements $x_1, \ldots x_r$ will be denoted by $(x_1, x_2, \ldots x_r)$ and the greatest number of linearly independent elements which can be chosen in the set will be called its order. This being so, let ϵ_2 be an element of F' which does not belong to A_1; then the set $A_2 = (1, \epsilon_1, \epsilon_2)$ is of order 3 since $\xi_0 + \xi_1 \epsilon_1 + \xi_2 \epsilon_2 = 0\ (\xi_2 \neq 0)$ gives $\epsilon_2 = -(\xi_0 + \xi_1 \epsilon_1)/\xi_2 < A_1$. We then form $A_3 = (1, \epsilon_1, \epsilon_2, \epsilon_3)$ by adding an element ϵ_3 of F' which is not in A_2; and so on. At some stage this process must come to an end, that is, for some integer i we have $A_{i-1} = (1, \epsilon_1, \ldots \epsilon_{i-1}) = F'$, since all the elements of F' also belong to $F(\alpha)$ whose basis is of order n. The product of any two ϵ's, being an element of F' since the latter is a field, must then be linearly dependent on the same set of ϵ's, that is, these form the basis of a linear associative algebra which is in fact the field F'. By Theorem 9.1 there is therefore an element β of F' such that $F' = F(\beta)$ i. e., F' is an algebraic field over F as required.

THEOREM 10.2. *If $F(\beta)$ is a subfield of $F(\alpha)$, the degree b of $F(\beta)$ is a factor of the degree a of $F(\alpha)$; and in $F(\beta)$ α satisfies an irreducible equation of degree a/b.*

If we denote elements of $F(\beta)$ by η with appropriate affixes, then, exactly as in the proof of the previous theorem, we can find a set of elements of $F(\alpha)$, $\epsilon_1 = 1, \epsilon_2, \ldots \epsilon_c$, such that (i) every element of $F(\alpha)$ has the form

$$y = \eta_1 \epsilon_1 + \eta_2 \epsilon_2 + \cdots \eta_c \epsilon_c \tag{10.1}$$

and (ii) there is no relation of the form $\eta_1 \epsilon_1 + \eta_2 \epsilon_2 + \cdots \eta_c \epsilon_c = 0$ except when all the η's are zero, i. e., the ϵ's are linearly independent with respect to $F(\beta)$.

Now if $\zeta_1, \zeta_2, \ldots \zeta_b$ is a basis of $F(\beta)$ and $\eta_i = \sum_j \xi_{ij} \zeta_j$, the ξ's as before being elements of F, then $y = \sum_{ij} \xi_{ij} \epsilon_i \zeta_j$. The elements $\epsilon_i \zeta_j$ ($i = 1, \ldots c$; $j = 1, \ldots b$) are linearly independent with respect to F since any linear relation among them would have the form

$$0 = \sum_i (\sum_j \xi_{ij} \zeta_j) \epsilon_i = \sum_i \eta_i \epsilon_i,$$

which can only be true if every $\eta_i = 0$, and this in turn is only possible if every $\xi_{ij} = 0$. It follows that $\epsilon_i \zeta_j$ ($i = 1, \ldots c$; $j = 1, \ldots b$) is a basis of $F(\alpha)$ and therefore $a = bc$ as required by the theorem.

From (10.1) it is evident that the elements of $F(\alpha)$ form a linear associative algebra of degree c with coefficients in $F(\beta)$. A variable element of this algebra satisfies an irreducible equation of degree c in $F(\beta)$ and this is also true of α seeing that it is a primitive element of the field.

11. **Normal fields.** If $\varphi(x) = 0$ is an irreducible equation in F of degree n we can evidently extend F by successive adjunction till we arrive at a field in which $\varphi(x)$ is resolvable into linear factors. Not more than $n - 1$ successive adjunctions will have to be made but the actual number of steps depends necessarily on the nature of $\varphi(x)$. If $\alpha_1, \alpha_2, \ldots \alpha_n$ are the roots of $\varphi(x)$ in the extended field, the latter is conveniently denoted by $F(\alpha_1, \alpha_2, \ldots \alpha_n)$ although it is not necessary to adjoin all the roots to obtain the full field. From Theorem 9.1 it follows that the same field can be obtained by a single adjunction, say $F(\varrho_1)$, where ϱ_1 is the root of an irreducible equation $g(x) = 0$ in F of degree μ, and as the α's belong to $F(\varrho_1)$, we may set

$$\alpha_i = \chi_i(\varrho_1), \qquad (i = 1, 2, \ldots n)$$

where $\chi_i(\varrho_1)$ is an integral function of ϱ_1 in F of degree $\mu - 1$ at most. Since $g(t)$ is irreducible, this is equivalent to saying that $\varphi(\chi_i(t)) \equiv 0 \bmod. g(t)$;

hence, if a second root ϱ_2 of $g(t) = 0$ is adjoined to $F(\varrho_1)$, we shall have $\varphi(\chi_i(\varrho_2)) = 0$ as well as $\varphi(\chi_i(\varrho_1)) = 0$ for $i = 1, 2, \ldots, n$. But φ cannot have more than n roots; hence $\chi_i(\varrho_2)$ is equal to one of the α's say α'_i, and the roots of φ are all integral functions of ϱ_2 in F. It follows that ϱ_1, being a polynomial in the α's, is also a polynomial in ϱ_2 or say

$$\varrho_1 = \theta(\varrho_2).$$

Hence $F(\varrho_1) \leq F(\varrho_2)$. The degree of $F(\varrho_2)$ cannot be greater than the degree μ of $F(\varrho_1)$ since both ϱ_1 and ϱ_2 satisfy the equation $g(t) = 0$ of degree μ, hence from Theorem 10.2 the degrees are equal and $F(\varrho_1) = F(\varrho_2)$. It follows that ϱ_2 is an element of $F(\varrho_1)$ and is therefore an integral function of ϱ_1 in F. Since ϱ_2 is any root of $g(t) = 0$, all the roots of this equation are rational integral functions of ϱ_1. Such an equation is called a *normal equation* and the corresponding field is called a *normal field*.

We have therefore proved

Theorem 11.1. *Every algebraic field over F is a subfield of a normal field.*

A THEOREM CONCERNING CERTAIN UNIT MATRICES WITH INTEGER ELEMENTS.

By H. R. Brahana.

In his *Cinquième Complément* to Analysis Situs* Poincaré makes the statement that the group of matrices, whose elements are integers and whose determinants are 1 and which leave the normal form of a skew-symmetric matrix of an even number of rows invariant, is generated by the matrices of the three following types of linear transformations on $2p$ variables:

$$(E_i) \qquad x'_{2i+1} = x_{2i+1} \pm x_{2i+2} \qquad (i = 0, 1, \ldots, p-1)$$

$$x'_j = x_j,$$

where j takes all values different from $2i + 1$.

$$(E''_i) \qquad x'_{2i+2} = x_{2i+2} \pm x_{2i+1} \qquad (i = 0, 1, \ldots, p-1)$$

$$x'_j = x_j,$$

where j takes all values different from $2i + 2$.

$$(E_{ij}) \qquad x'_{2i+1} = x_{2i+1} \pm x_{2j+1} \qquad (j \neq i = 0, 1, 2, \ldots, p-1)$$

$$x'_{2j+2} = x_{2j+2} \pm x_{2i+2}$$

$$x'_k = x_k,$$

where k takes all values different from $2i + 1$ and $2j + 2$.

The present writer has been unable to find a proof of this statement in the literature and having use for the fact stated has been led to establish it. It is hoped that the theorem and its proof may be of interest in themselves.

1. **Statement of the problem.** If M is the skew-symmetric square matrix of an even number of rows

* Rendiconti del Cir. Mat. di Palermo, vol. 18, p. 65.

$$\begin{Vmatrix} 0 & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ -1 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & 0 & 1 & \cdot & \cdot & \cdot & 0 & 0 \\ 0 & 0 & -1 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 1 \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot & -1 & 0 \end{Vmatrix}$$

and A is a square matrix whose elements are integers and whose determinant is 1, we have the following theorem.

If A satisfies the condition $AMA' = M$, then A is a product of matrices of the types (E_i), $E'_j)$ and (E_{ij}).

2. **Reduction of the problem.** Let us consider the relation $AMA' = M$. Multiplying both sides of the equation on the right by $(A')^{-1}$ and on the left by M we get a relation which we may use to find A^{-1}, viz. $MAM = -(A')^{-1} = -(A^{-1})'$. Thus, if

$$A = \begin{Vmatrix} a_{11} & b_{11} & a_{12} & b_{12} & \cdot & \cdot & \cdot & a_{1p} & b_{1p} \\ c_{11} & d_{11} & c_{12} & d_{12} & \cdot & \cdot & \cdot & c_{1p} & d_{1p} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{p1} & b_{p1} & a_{p2} & b_{p2} & \cdot & \cdot & \cdot & a_{pp} & b_{pp} \\ c_{p1} & d_{p1} & c_{p2} & d_{p2} & \cdot & \cdot & \cdot & c_{pp} & d_{pp} \end{Vmatrix},$$

we get

$$A^{-1} = \begin{Vmatrix} d_{11} & -b_{11} & d_{21} & -b_{21} & \cdot & \cdot & \cdot & d_{p1} & -b_{p1} \\ -c_{11} & a_{11} & -c_{21} & a_{21} & \cdot & \cdot & \cdot & -c_{p1} & a_{p1} \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ d_{1p} & -b_{1p} & d_{2p} & -b_{2p} & \cdot & \cdot & \cdot & d_{pp} & -b_{pp} \\ -c_{1p} & a_{1p} & -c_{2p} & a_{2p} & \cdot & \cdot & \cdot & -c_{pp} & a_{pp} \end{Vmatrix}.$$

The element c_{11} of the product $A \cdot A^{-1}$ is $a_{11}\, d_{11} - b_{11}\, c_{11} + a_{12}\, d_{12} - b_{12}\, c_{12} + \cdots + a_{pp}\, d_{pp} - b_{pp}\, c_{pp}$. This element must be 1; therefore we get the result that the sum of the determinants of the two-rowed square matrices obtained by taking the elements of the first and second columns, the third and fourth columns, and so on up to the $(2p-1)$ and $2p$th columns in the first and second rows of A is 1. In the same manner the sum of determinants of corresponding matrices in the third and fourth rows is 1.

It will be convenient for our purposes to consider the transformations in the following manner. Transformations (E_i) and (E_i'') are equivalent respectively to adding any odd row to the succeeding row, and adding any even row to the preceding row (or the inverse operations when the negative sign is used). Transformation (E_{ij}) is equivalent to adding any odd row to any other odd row, provided that the corresponding even rows are subtracted in the reverse order; thus we may add the first row to the third provided we subtract the fourth row from the second, or we may subtract the fifth row from the first provided we add the second row to the sixth. If the matrices of the transformations are multiplied in on the right, the operations just described are performed upon columns instead of upon rows. Regarding the transformations from this point of view it is clear that the inverse of any one of types (E_i), (E_i''), and (E_{ij}) is a transformation of the same type. Since these operations leave M invariant, the product B of A and any number of these matrices on both the left and the right is such that $BMB' = M$.

We will now see that by means of these operations any unit matrix A with integer elements may be reduced to a form in which every element above the main diagonal is 0, and in which each of the two-rowed minors on the main diagonal and in the first and second, third and fourth, ... $(2p-1)$th and $2p$th columns is $\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}$.

Since the matrix A is of determinant 1 the elements of the first row are relatively prime. We will combine the first and second columns in such a way as to reduce the element a_{12} to zero. By performing transformations (E_i) and (E_i'') on the other columns in like manner, we may reduce each of the two rowed matrices in the first two rows of A to the same form. If a_{11} is not 1, transformation (E_{ij}) may be applied in columns in the following manner. First obtain an element a_{11} which is not zero by adding some column to the first, then obtain an element a_{1i} a positive number smaller than a_{11}; this may be done by subtracting the first column from the ith column a sufficient number of times, and adding the $(i+1)$th column to the second the same number of times. This accomplished, subtract the ith column from the first and add the second to the $(i+1)$th. By repetition of this process we may obtain an a_{11} which is equal to 1.

The two-rowed matrices have not been disturbed by the above operations. They are still in the form $\begin{Vmatrix} a_{1i} & 0 \\ c_{1i} & d_{1i} \end{Vmatrix}$. Then by subtracting the first column from each odd column a sufficient number of times we may obtain a first row which is made up of zero elements with the exception of a_{11}. Recalling the fact that the sum of the determinants in the first two rows is 1 and noting that the first one is the only one which is not zero we see that the two-rowed square matrix in the upper left-hand corner is of determinant 1, and that the $(2p-2)$-rowed square matrix in the lower right-hand corner has a determinant whose value is the same. The two-rowed matrix in the upper left-hand corner can be put in the form $\begin{Vmatrix} 1 & 0 \\ 0 & 1 \end{Vmatrix}$ by operations on the first and second columns.

Now each of the two-rowed minors in the first two rows (with the exception of the one on the main diagonal) is in the form $\begin{Vmatrix} 0 & 0 \\ c_{1i} & d_{1i} \end{Vmatrix}$. By means of transformations (E_i) and (E'_i) each of these may be put in the form $\begin{Vmatrix} 0 & 0 \\ 0 & d_{1i} \end{Vmatrix}$. Then, applying transformation (E_{ij}) in such a way as to add the second column to the $2i$th column $-d_{1i}$ times, we may reduce each of the two-rowed minors to $\begin{Vmatrix} 0 & 0 \\ 0 & 0 \end{Vmatrix}$. This does not affect the two-rowed matrix in the upper left-hand corner.

The $(2p-2)$-rowed square matrix in the lower right-hand corner is of the same sort as the original matrix of $2p$ rows and so may be reduced by transformations (E_i), (E'_i), and (E_{ij}) to the desired form.

3. **Solution of the problem.** Let A be of the form

$$\begin{Vmatrix}
1 & 0 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\
0 & 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\
a_{21} & b_{21} & 1 & 0 & \cdot & \cdot & \cdot & 0 & 0 \\
c_{21} & d_{21} & 0 & 1 & \cdot & \cdot & \cdot & 0 & 0 \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
a_{p1} & b_{p1} & a_{p2} & b_{p2} & \cdot & \cdot & \cdot & 1 & 0 \\
c_{p1} & d_{p1} & c_{p2} & d_{p2} & \cdot & \cdot & \cdot & 0 & 1
\end{Vmatrix}$$

and let us now consider the product AMA'.

It will be skew-symmetric. If we impose the condition that $AMA' = M$ we obtain a system of $[(2p)^2 - 2p]\,2$ equations in $2p(p-1)$ unknowns.

The equations are not all independent; they reduce to $2p(p-1)$ equations. There is a unique solution of the system; each of the elements below the main diagonal is zero. This may be seen very easily in another way. Consider the product $AMA' = M$. The first two rows of the product AM will be the same as the first two rows of M. Then the first and second rows of AMA' will be respectively the second and the negative of the first rows of A'; viz.

$$\begin{matrix} 0 & 1 & b_{21} & d_{21} & \cdot & \cdot & \cdot & b_{p1} & d_{p1} \\ -1 & 0 & -a_{21} & -c_{21} & \cdot & \cdot & \cdot & -a_{p1} & -c_{p1}. \end{matrix}$$

This product, however, must be equal to M. Thus we see that each of the unknown elements in the first two columns of A must be zero. Continuing in this manner we determine each of the unknown elements of A. Since A is the identity matrix, the original matrix must have been a product of types (E_i), (E_i') and (E_{ij}).

This completes the proof of the theorem.

4. **Normalization of any skew-symmetric matrix which is of determinant 1, and which has integer elements.** Given any skew-symmetric matrix N with integer elements and of determinant 1, we see, as a result of the theorem just proved, that there are an infinite number of matrices C which satisfy the relation $CNC' = M$, (M is the normal form). For example, if B is a particular matrix satisfying the relation $BNB' = M$ then the product C of B multiplied on the left by any number of matrices of types (E_i), (E_i') and (E_{ij}) also satisfies the relation $CNC' = M$. We may now consider the following question. Does there exist any matrix C, satisfying the relation $CNC' = M$, which cannot be obtained by multiplying a particular matrix B on the left by matrices of types (E_i), (E_i') and (E_{ij})? The answer is that there is not. Let B be a particular matrix which satisfies the relation

$$BNB' = M.$$

then

$$N = B^{-1}M(B')^{-1}.$$

If C is any other matrix satisfying the relation

$$CNC' = M.$$

then

$$CB^{-1}M(B')^{-1}C' = M$$

and

$$CB^{-1} = A,$$

whence

$$C = AB.$$

Thus, if C is any matrix which normalizes N, it may be obtained by taking any other matrix satisfying that relation and multiplying it on the left by matrices of types (E_i), (E_i') and (E_{ij}).

Now consider any matrix D which satisfies the relation

$$DND' = N.$$

Then, if B is the particular matrix described above, we get the relation

$$BDND'B' = BNB' = M.$$

The product BD is of the type C above. Then we have $BD = AB$, where A is a matrix made up of products of types (E_i), (E_i') and (E_{ij}). Hence

$$D = B^{-1}AB.$$

The result may be stated in the following form: *The group of matrices which satisfy the relation $DND' = N$ is the transform of the group of matrices A which satisfy the relation $AMA' = M$ by any matrix B which satisfies the relation $BNB' = M$.*

AN INTRODUCTION TO THE THEORY OF ELLIPTIC FUNCTIONS.*

By Gösta Mittag-Leffler.

Preface to the English edition.

When my paper was first published it was my intention to let it form a part of a greater memoir, comprehending a detailed exposition and critical analysis of all the different methods which form an introduction to the theory of elliptic functions. However, my time became fully engaged by other scientific occupations so I could never prevail upon myself to take up and carry through my original intention. Since I hoped to be able to carry through my original plan, my paper was never translated into a foreign language. I have found, however, that I cannot expect my hope to be fulfilled; on the other hand, the passing years have not seen the appearance of a satisfactory analysis of connections between Abel and Weierstrass, which, by the way, extends to many other fields than that of elliptic functions. Under these circumstances I have accepted with sincere gratitude the offer made by Professor E. Hille to translate my memoir into English and I wish to offer him my thanks for the energy and knowledge of the subject matter which he has expended on the present paper.

Preface to the Swedish edition.†

> Ich fordere, man soll bei allem Gebrauch des Kalküls, bei allen Begriffsverwendungen sich immer der ursprünglichen Bedingungen bewußt bleiben und alle Produkte des Mechanismus niemals über die klare Befugnis hinaus als Eigentum betrachten.
>
> Gauss.

Half a century has not yet elapsed since the elliptic functions were first introduced in mathematics. From that time on the theory has increased to such an extent that nowadays scarcely any other field of mathematics can offer such an abundance of formal results and such a wealth of applications to

* Authorized translation from the Swedish by Einar Hille.

† [The present paper was published as a separate pamphlet in Hälsingfors in March 1876 under the title "En metod att komma i analytisk besittning af de Elliptiska Funktionerna". It was presented as an academic dissertation by Professor Mittag-Leffler to the then Imperial Alexander University in Finland when competing for the chair in mathematics in that university which he subsequently held from 1877 to 1881. Translator's note.]

different branches of the exact sciences. Moreover, the prophetic divination of Euler has become a reality, *the discovery of this theory has essentially extended the bounds of mathematical analysis.* New fields have been opened for mathematical thought and the number of fundamental ideas with which mathematics operates has been vastly increased. A careful analysis of these fundamental ideas has formed the point of departure of a great number of the investigations, the results of which form the peaks of our present day knowledge in mathematics. This source of new and essential progress is certainly far from exhausted.

However, while mathematical literature is rich in memoirs on elliptic functions, we still lack a comparative exposition and analysis of the different methods which offer an introduction to the theory of elliptic functions, and of the fundamental ideas which form the starting-points of these methods and which thereby constitute them as essentially different methods. At the suggestion of Weierstrass the author of this memoir has for some time past been occupied with the solution of *the problem to determine by what essentially different methods an introduction to the theory of elliptic functions can be obtained.*

The present paper forms a part of the author's investigation which can be considered as closed in itself. Its purpose is, on one hand, to expound the fundamental ideas which prepared the way and made a theory of elliptic functions possible, on the other hand, to give a rigorous presentation of one of the main roads which leads to an introduction to this theory. This road, the road of Abel, is historically the oldest, and it is also founded on ideas which belong to the lower layers of the mathematical structure, thus leading to the goal with a minimum number of assumptions. If the number of assumptions is increased, the method of Abel gains in unity and perfection and passes over into the older method of Weierstrass in a natural and simple manner. Thus a description of the latter method belongs to the field of the problem whose solution we aim at in the present paper. It also belongs to this field to show the deficiencies of Abel's method as well as that of Weierstrass and to indicate how these can be remedied by other methods.

Introduction.

Once in possession of those elementary integrals which are algebraic functions, inverse trigonometric functions and logarithms, mathematicians naturally tried reducing as many problems as possible in integration to such integrals. It was soon found that every endeavor to carry through such a reduction by means of a finite number of algebraic operations was

in vain in many cases. This was especially the case with the integrals contained in the general formula

$$(1) \qquad \int^{x} \frac{F(x)}{\sqrt{R(x)}}\, dx,$$

where $F(x)$ is a polynomial in x and

$$(2) \qquad R(x) = Ax^4 + 4Bx^3 + 6Cx^2 + 4B'x + A'.$$

When either or both of the coefficients A and B had values different from zero and the roots of $R(x) = 0$ were all unequal, it proved impossible to reduce the integral (1) to the known integrals mentioned above.

The treatment of this integral was not a problem arbitrarily proposed, but belonged to the class of problems which are put by mathematics itself so to speak. In fact, the integrals which express the length of arc of an ellipse and of an hyperbola fall under the general formula (1). And from the geometrical point of view, which prevailed in higher analysis during its earlier stage, the problem of rectifying the ellipse and the hyperbola would be one of the first problems to attract the attention of mathematicians.

When every effort to reduce the integral (1) to known elementary forms of integrals had failed, nothing was left for mathematicians but to make up their minds to introduce new transcendental functions in analysis. Two roads were open for investigation. On the one hand, one could solve the formal problem of reducing the number of new transcendents to a minimum; on the other hand, one had to investigate what might be called the intrinsic properties of these transcendental functions which it was necessary to incorporate with analysis. Many investigations were carried through along the lines first mentioned* without yielding any more important results however. In the main they tried to reduce as many as possible of the integrals contained in the general formula (1), to the particular integral that represents the length of arc of an ellipse. The geometrical point of view, considered indispensible in these days, led to the erroneous assumption that this integral is the simplest of all integrals

* MacLaurin: A Treatise of Fluxions. (1742.) Vincenti Riccati: Opuscula. (1742.) D'Alembert, "Mémoires de l'académie des sciences et des belles lettres". Berlin. (1746, 1748.) Euler also treated this question in several papers published by the Academy of Sciences in Petersburg. This type of investigation was brought to a finish by Legendre when he introduced the three normal forms for the elliptic integrals.

of the form (1). This is also the reason why the integrals of this form have received the name *elliptic integrals** which is rather inappropriate.

The latter problem, to investigate the intrinsic properties of these elliptic functions, was a much harder one in view of the imperfect understanding of the spirit of analysis which then prevailed.

However, one fact had early attracted attention and indicated the direction in which the problem had to be attacked. It was known that the differential equation

$$f(x)\,dx = \pm f(y)\,dy,$$

where $\int f(x)\,dx$ is a logarithm or an inverse trigonometric function, had as an integral an *algebraic function* of x and y and consequently, *in spite of the fact that it is impossible to find an algebraic integral of the differential $f(x)\,dx$, one can find an algebraic integral of the sum or the difference of two such differentials.*

Then it was natural to ask whether this property might hold for other transcendents than the logarithm and the inverse trigonometric functions. This question is due chiefly to Johannis Bernoulli.†

The first answer to this question which contained an essentially new contribution was given by Fagnano.‡ He had already, in 1715, proved the following theorem:

If in the expression

$$(A)\qquad \frac{x^{n-1}(x^n+p)^{h-1}\,dx}{[(x^n+p)^2+q(x^n+p)+r]^h}$$

where p, q and r are arbitrary constants and n and h rational numbers, a new variable z is introduced by the relation

$$z^n x^n + p(z^n + x^n) + p^2 = r,$$

---

* Legendre called them *elliptic functions*. Jacobi, however, gave this name to the inverse functions, and since then the name *elliptic integrals* for the integrals of type (1) has been recognized in mathematics. Cf. for this question "Lettre de Legendre à Jacobi: Réponse de Jacobi", Journal für Mathematik, vol. 80 (1875), pp. 269-270.

† Cf. a paper by Johannis Bernoulli in Acta Erudit. Lips. (1698), p. 462. Also "Addition historique à la préface d'un mémoire de Lagrange par Jean Plana", Memoria della reale accademia della scienze di Torino. Ser. 2, vol. 20 and the introduction of "Nouvo metodo . . .", Produzione Matematiche Del Marchese Giulio Carlo De' Toschi Di Fagnano, vol. 2, p. 317.

‡ Loc. cit. in Produzione Matematiche, etc.

then (A) is transformed into

$$-\frac{z^{n-1}(z^n+p)^{h-1}\,dz}{[(z^n+p)^2+q(z^n+p)+r]^h}.$$

If we put $n=2$, $h=\frac{1}{2}$ and either $p=0$ or $p^2+pq+r=0$, we obtain as a special case of Fagnano's theorem: *a differential equation of the form*

$$\frac{dx}{\sqrt{f+gx^2+hx^4}}+\frac{dz}{\sqrt{f+gz^2+hz^4}}=0,$$

is satisfied by an algebraic function of x and z which is symmetric of degree two in each of these variables.

Fagnano proved several similar theorems in later papers* which also concern certain types of elliptic integrals. Thus Fagnano had given an answer to the question of Bernoulli which gave real new insight into the theory of transcendental functions. He had shown that the same remarkable algebraic property which characterized the logarithm and the inverse trigonometric functions held for at least certain classes of elliptic integrals.

The theorems of Fagnano gave the first deeper insight into the nature of the elliptic integrals and hence formed the starting point for the further progress toward the discovery of the principal properties of these functions. As a matter of fact, Euler takes Fagnano's results as his point of departure in the famous memoir,† *De integratione aequationis differentialis*

$$\frac{m\,dx}{\sqrt{1-x^4}}=\frac{n\,dy}{\sqrt{1-y^4}}. \tag{3}$$

The fundamental criticism that could be raised against Fagnano's results was that the algebraic integrals he obtained of the differential equations in question, all were *particular integrals*. Now it is obvious that, even if a differential equation has a *particular algebraic integral*, the *complete integral* might very well be a *transcendental function*. For this reason Euler took upon himself the task of finding an algebraic form for the complete integral of the differential equation

* See Produzione Matematiche etc. Further cf. the note of F. Siacci, "Sul Teorema del Conte di Fagnano", Bulletino di Bibliografia e di Storia della Scienze Matematiche e Fisiche, pubblicato da B. Boncompagni, vol. 3 (1870).

† Novi Commentarii Acad. Scient. Petropolitanae, vol. 6 (1756–57), p. 37.

$$\frac{dx}{\sqrt{1-x^4}} = \frac{dy}{\sqrt{1-y^4}}, \tag{4}$$

and of deducing from this an algebraic form of the complete integral of the more general equation (3).

From the investigations of Fagnano it follows that equation (4) is satisfied by the particular integral

$$x = -\sqrt{\frac{1-y^2}{1+y^2}}.$$

Another particular integral is evidently

$$x = y.$$

Thus the complete integral must be such that it reduces to one or the other of these two particular integrals for special values of the arbitrary constant. Guided by the form of these integrals, Euler finds, "*potius tentando vel divinando*" that the complete integral is given by

$$x = \frac{y\sqrt{1-c^4}+c\sqrt{1-y^4}}{1+c^2y^2}, \tag{5}$$

where c is an arbitrary constant.

To show that (5) really is the desired integral, one has only to differentiate and to eliminate c.

Already in the memoir in question, Euler draws several conclusions of fundamental importance from the form in which the integral (5) appears. One can even claim that the fundamental ideas concerning *the addition and multiplication of elliptic integrals of the first order* which were known before the work of Abel, are actually developed in this memoir by Euler. This does not seem to be so well known* as might be expected, and the reason is probably in the main the peculiar geometric form in which Euler considered himself bound to express his ideas. This form is, however, not at all necessary and we now proceed to show briefly how Euler's leading ideas can actually be expressed analytically in detail.

* Cf. Poisson, "Rapport sur l'ouvrage de M. Jacobi, intitulé Fondamenta nova etc.", Mémoires de l'académie royale des sciences de l'institute de France, vol. 10, p. 76.

(5) is a complete integral of (4), but such is also the case with

$$(6) \qquad \int_0^x \frac{dx}{\sqrt{1-x^4}} = \int_0^y \frac{dx}{\sqrt{1-x^4}} + \int_0^k \frac{dx}{\sqrt{1-x^4}},$$

where k is an arbitrary constant. This equation has a definite meaning, at least when the variable of integration, x, is always real and the three numbers x, y and k lie between -1 and $+1$. If one takes for granted that a differential equation in two variables can have only *one complete* integral, then the equations (5) and (6) are to be considered as different analytic expressions for the same functional relation between x and y, if only k is appropriately determined in terms of c. If we put $y = 0$ in (5), we get $x = c$; and if these values of y and x are substituted in (6), we obtain $k = c$. Hence, *if the sum of two elliptic integrals of the same form and the same lower limits*

$$\int_0^y \frac{dx}{\sqrt{1-x^4}} + \int_0^c \frac{dx}{\sqrt{1-x^4}}$$

is equated to a third integral, also of the same form and with the same lower limit, viz.,

$$\int_0^x \frac{dx}{\sqrt{1-x^4}} = \int_0^y \frac{dx}{\sqrt{1-x^4}} + \int_0^c \frac{dx}{\sqrt{1-x^4}},$$

then the upper limit x of the latter integral is an algebraic symmetric function (5) *of the arbitrarily chosen upper limits, y and c, of the other two integrals.*

This is the so-called *addition theorem* for the elliptic integral

$$\int_0^x \frac{dx}{\sqrt{1-x^4}}.$$

This theorem is often called *Euler's addition theorem* because the addition theorem coincides with the preceding theorem according to which the complete integral of equation (4) is a rational entire function of x and y and the arbitrary constant c of form (5).

If we put $c = y$ in (7), then (5) and (7) still give the same value of x for the same value of y. We designate this value by y_1 and obtain

$$(8)\quad \begin{aligned} &(a) \qquad y_1 = \frac{2y\sqrt{1-y^4}}{1+y^4}, \\ &(b) \qquad \int_0^{y_1} \frac{dx}{\sqrt{1-x^4}} = 2\int_0^{y} \frac{dx}{\sqrt{1-x^4}}. \end{aligned}$$

Then we put $c = y_1$ and designate the corresponding value of x by y_2, obtaining

$$(9)\quad \begin{aligned} &(a) \qquad y_2 = \frac{y\sqrt{1-y_1^4}+y_1\sqrt{1-y^4}}{1+y_1^2\,y^2}, \\ &(b) \qquad \int_0^{y_2} \frac{dx}{\sqrt{1-x^4}} = 3\int_0^{y} \frac{dx}{\sqrt{1-x^4}}. \end{aligned}$$

If this process is repeated by substituting instead of c in (5) and (7) successively $y, y_1, y_2, \ldots y_{n-2}$, then y_1 is determined by (8a), y_2 by (9a) and so on; thus we obtain a sequence of $(n-1)$ equations of the same form as (8) and (9). The last these equations is

$$(10)\quad \begin{aligned} &(a) \qquad y_{n-1} = \frac{y\sqrt{1-y_{n-2}^4}+y_{n-2}\sqrt{1-y^4}}{1+y_{n-2}^2\,y^2}, \\ &(b) \qquad \int_0^{y_{n-1}} \frac{dx}{\sqrt{1-x^4}} = n\int_0^{y} \frac{dx}{\sqrt{1-x^4}}. \end{aligned}$$

Thus, if the n^{th} multiple of the elliptic integral

$$\int_0^{x} \frac{dx}{\sqrt{1-x^4}}$$

is equated to another elliptic integral of the same form and with the same lower limit, then the upper limit of the latter integral is a certain algebraic function of the upper limit of the first integral. This is the so-called *multiplication theorem* for the elliptic integral

$$\int_0^x \frac{dx}{\sqrt{1-x^4}}.$$

It is now easy to represent in algebraic form the complete integral of the equation

$$\frac{dy_n}{\sqrt{1-y_n^4}} = n\frac{dy}{\sqrt{1-y^4}}. \tag{11}$$

The transcendental form of the complete integral is

$$\int_0^{y_n} \frac{dx}{\sqrt{1-x^4}} = n\int_0^y \frac{dx}{\sqrt{1-x^4}} + \int_0^c \frac{dx}{\sqrt{1-x^4}}. \tag{12}$$

In view of the theorems of addition and of multiplication, we have the following relation between the upper limits in equation (12), namely

$$y_n = \frac{y_{n-1}\sqrt{1-c^4} + c\sqrt{1-y_{n-1}^4}}{1+c^2 y_{n-1}^2}, \tag{13}$$

and this is the desired algebraic expression for the complete integral of (11). The complete integral of (3) is then immediately obtainable. On account of the multiplication theorem, the equation

$$\frac{dx_{m-1}}{\sqrt{1-x_{m-1}^4}} = m\frac{dx}{\sqrt{1-x^4}} \tag{14}$$

admits of an integral of the form

$$x_{m-1} = \frac{x\sqrt{1-x_{m-2}^4} + x_{m-2}\sqrt{1-x^4}}{1+x_{m-2}^2 x^2}, \tag{15}$$

where the meaning of x_{m-2}, x_{m-3} etc. is clear from the preceding text. If we introduce y_n instead of x_{m-1} in (14) and (15) we find, comparing with (13), that the complete integral of equation (3) is

$$(16)\qquad \frac{x\sqrt{1-x_{m-2}^4}+x_{m-2}\sqrt{1-x^4}}{1+x_{m-2}^2\,x^2}=\frac{y_{n-1}\sqrt{1-c^4}+c\sqrt{1-y_{n-1}^4}}{1+c^2\,y_{n-1}^2},$$

where the meaning of x_{m-2}, x_{m-3}, etc., y_{n-1}, y_{n-2}, etc. is obvious from the text above.

Thus we have found that *the complete integral of the differential equation* (3) *is expressible as an algebraic function of the two variables x and y and an arbitrary constant.*

Euler ends the memoir in question by showing how the theorems of addition and multiplication, as well as the result just obtained, can be extended to hold for a differential equation of the general form

$$(17)\qquad \frac{m\,dx}{\sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}}=\frac{n\,dy}{\sqrt{Ay^4+6By^3+6Cy^2+4B'y+A'}}.$$

Euler himself saw a great defect in his investigations in the fact that he had not succeeded in obtaining his complete algebraic integral by a general method of analysis and that consequently his results were without natural connection with other parts of the Calculus. His challenge to mathematicians* to improve on this flaw was answered by Lagrange in 1768.†

Lagrange himself designates the following observation as the idea permeating his investigation. If a differential equation of the first order is given, of which the solution is unknown, we may differentiate the equation and investigate whether, by combining the new equation with the old one, we can obtain a complete integral of the first order which does not embrace as a particular case the original differential equation. If this process works, then the desired complete integral is obtained as the result of eliminating the derivative between the two equations of the first order.

Lagrange now undertakes to apply this principle to the problem of finding the complete integral of the differential equation

* Loc. cit. p. 20.

† "Sur l'intégration de quelques équations différentielles dont les indéterminées sont séparées mais dont chaque membre en particulier n'est point intégrable", Miscellanea Tauriensia, vol. 4 (1766–69), and Œuvres de Lagrange, vol. 2.

$$\frac{dx}{\sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}} \tag{18}$$
$$= \frac{dy}{\sqrt{Ay^4+4By^3+6Cy^2+4B'y+A'}}.$$

Instead of this equation he introduces the system of simultaneous equations

$$\begin{aligned} \frac{dx}{dt} &= \sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}, \\ \frac{dy}{dt} &= \sqrt{Ay^4+4By^3+6Cy^2+4B'y+A'}, \end{aligned} \tag{19}$$

and then replaces the variables x and y by two new variables p and q determined by the equation

$$x+y=p, \qquad x-y=q.$$

If we put

$$\begin{aligned} Ax^4+4Bx^3+6Cx^2+4B'x+A' &= X, \\ Ay^4+4By^3+6Cy^2+4B'y+A' &= Y, \end{aligned} \tag{20}$$

we obtain

$$\frac{dp}{dt} = \sqrt{X}+\sqrt{Y}, \qquad \frac{dq}{dt} = \sqrt{X}-\sqrt{Y}. \tag{21}$$

Differentiating the first equation, we get

$$\frac{d^2p}{dt^2} = \frac{1}{2}\left(\frac{dX}{dx}+\frac{dY}{dy}\right). \tag{22}$$

It remains to introduce p and q in (21) and (22) instead of the old variables, and then to combine these equations in such a way that a new differential equation of order one is obtained.

By calculations of great elegance Lagrange forms out of the preceding equations the equality

$$\frac{d}{dt}\left[\left(\frac{1}{q}\,\frac{dp}{dt}\right)^2\right] = 2Ap\frac{dp}{dt}+4B\frac{dp}{dt}, \tag{23}$$

which yields by integration

$$\left(\frac{1}{q}\frac{dp}{dt}\right)^2 = Ap^2 + 4Bp + G. \tag{24}$$

where G is an arbitrary constant.

If this equation is combined with the first equation in (21) and p and q are expressed in terms of x and y, we find that the complete integral of (18) is given by

$$\begin{aligned}\sqrt{Ax^4 + 4Bx^3 + 6Cx^2 + 4B'x + A'}\\ + \sqrt{Ay^4 + 4By^3 + 6Cy^2 + 4B'y + A'}\\ = (x-y)\sqrt{A(x+y)^2 + 4B(x+y) + G}.\end{aligned} \tag{25}$$

Lagrange tried in vain to give his method a more general application and to use it for the integration of similar differential equations where the algebraic function under the radical was of higher degree than the fourth. Euler, also, tried without success to extend the validity of his addition theorem in that direction.

It remained for Abel to present the solution of this problem which is called *Abel's general addition theorem*. To use the words of Legendre, this product of the deepest mathematical thinking has raised for its originator a "*monumentum aere perennius*". Just as *Euler's addition theorem*, as we shall soon see, formed the foundation upon which Abel and Jacobi erected *the theory of elliptic functions*, *Abel's addition theorem* later became the base upon which Weierstrass and Riemann built the new and magnificent theory which on Jacobi's suggestion received the name of the *theory of Abelian functions*.

Richelot showed much later how the method of Lagrange applies to the solution of a much more general problem of addition than that of Euler.* Before the publication of the memoir of Lagrange, as early as in volume 7 of Novi Commentarii Academiæ Scientiarum Petropolitanæ,† Euler had advanced an analytical method for the deduction of the addition theorem which equals the method of Lagrange in elegance, and which leads in a very natural manner to the desired result if once grasped.‡

* "Über die Integration eines merkwürdigen Systems Differentialgleichungen", Journal für Mathematik, vol. 23, (1842), p. 354.

† Demonstratio theorematis et solutio problematis in Actis Erud. Lipsiensibus propositorum, p. 129.

‡ The deduction of the addition theorem according to Euler, is to be found in vol. 6 of Novi Commentarii.

Instead of trying to find an algebraic form of the complete integral of (18), one could set the problem of trying to transform the differential

$$\frac{dx}{\sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}} \tag{26}$$

by an algebraic transformation into another differential of the same form and with the same constants

$$\frac{dy}{\sqrt{Ay^4+4By^3+6Cy^2+4B'y+A'}} \tag{27}$$

or, in other words, the problem of expressing x as such an algebraic function of a new variable y that the differential (26) is carried over into the differential (27) by this transformation.

If such an algebraic equation should exist through which the differential (26) is transformed into the differential (27), with the sign plus or the sign minus, and if, moreover, this equation should contain an *arbitrary constant*, then that expression would yield the complete integral of the differential equation

$$\begin{aligned}\frac{dx}{\sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}}\\ = \pm\frac{dy}{\sqrt{Ay^4+4By^3+6Cy^2+4B'y+A'}}.\end{aligned} \tag{28}$$

In order to arrive at the solution of this equation in a systematic manner, one might first investigate whether the differential (26) can be transformed into the differential (27) by an algebraic, rational, and integral relation between the variables x and y which is of degree one in both of these variables. It is easy to show that this is impossible, except for the trivial transformation $x=y$; the reason being that such a relation contains only *three* essential constants, whereas the differential in question contains *five*. The next step is then to investigate whether the desired transformation can be effected by a rational integral relation between x and y of the *second* degree in each variable.

The most general form of such an equation is

$$\begin{aligned}(ax^2+2bx+c)y^2+2(a'x^2+2b'x+c')y+a''x^2+2b''x+c''\\ =(ay^2+2a'y+a'')x^2+2(by^2+2b'y+b'')x+cy^2+2c'y+c''=0.\end{aligned} \tag{29}$$

From this it follows immediately that

$$(ax^2+2bx+c)y+(a'x^2+2b'x+c')$$
$$=\sqrt{(a'x^2+2b'x+c')^2-(ax^2+2bx+c)(a''x^2+2b''x+c'')},$$
$$(ay^2+2a'y+a'')x+(by^2+2b'y+b'')$$
$$=\sqrt{(by^2+2b'y+b'')^2-(ay^2+2a'y+a'')(cy^2+2c'y+c'')}.$$

The ambiguity in sign of the radical is removed if we determine arbitrarily for one of the radicals which of the two possible values we give it at a certain point; the meaning of the second radical is then also uniquely determined.

By differentiation we get further

$$\{(ay^2+2a'y+a'')x+(by^2+2b'y+b'')\}\,dx$$
$$+\{(ax^2+2bx+c)\,y+(a'x^2+2b'x+c')\}\,dy=0,$$

thus obtaining

$$\frac{dx}{\sqrt{(a'x^2+2b'x+c')^2-(ax^2+2bx+c)(a''x^2+2b''x+c'')}}$$
$$(30)\qquad +\frac{dy}{\sqrt{(by^2+2b'y+b'')^2-(ay^2+2a'y+a'')(cy^2+2c'y+c'')}}=0.$$

If we consider the way in which the two functions under the radicals are formed, we notice that they become identical* if we put

$$(31)\qquad a'=b,\quad a''=c,\quad b''=c'.$$

The equation (30) then becomes

$$\frac{dx}{\sqrt{(bx^2+2b'x+c')^2-(ax^2+2bx+c)(cx^2+2c'x+c'')}}$$
$$(32)\qquad +\frac{dy}{\sqrt{(by^2+2b'y+c')^2-(ay^2+2by+c)(cy^2+2c'y+c'')}}=0.$$

Thus we have introduced *three* algebraic conditional equations (31) between the 9 constants contained in (29). If we further introduce the *five* conditional equations which are necessary in order that the differential equation (32)

* Euler does not set up the general equation (29) but starts from an equation of a more special form in which the constants are already subject to three conditions. Cf. Jacobi. Math. Werke, vol. 3, p. 85.

shall coincide with equation (28), we have *eight* independent algebraic equations between the *nine* constants in the transformation (29). We have considered only the upper sign in (28); it is easy to see how the method has to be modified if the lower sign is also taken into consideration. Thus, *eight* of the constants in (29) are expressible in terms of the *ninth* one and, consequently, the transformation (29) contains *one* arbitrary constant, thereby yielding the complete integral of (28).

We omit the calculation by which this complete integral is formed and restrict ourselves merely to an indication of Euler's leading idea as given above. We shall develop this idea further in § 1 below.

The transformation (29) is of the greatest importance even in other respects, and out of it follows in the most natural manner, not only the addition theorem, but also the other fundamental theorems of which the theory of elliptic functions was in possession before the time of Abel and Jacobi.

From this relation Euler deduced the addition theorem for the general elliptic integral (1) but, for the purpose at which we are aiming in this memoir, we do not need the addition theorem in any form more general than the one already developed and hence we shall not spend any more time on this consideration.

The differential equation (30) is an immediate consequence of the general transformation equation (29). In (29) nine independent constants enter and all of these occur in the first term of (30). Consequently, the latter expression can be made identical with the differential in (26), and that in infinitely many ways since we have *four* arbitrary constants at our disposal. We have seen already how *three* of these constants can be disposed of in such a way as to carry the differential in (26) over into the differential in (27).

The question presents itself whether it is possible to dispose of the arbitrary constants in such a way that the differential (26) is not transformed into itself but into a similar expression of simpler form. Such a form would be, for instance, one which contained only even powers of the variable. A glance at (30) shows that such a transformation is obtained by putting

$$a' = b' = c' = 0, \tag{33}$$

and $-y$ instead of y.

The equation (30) is then carried over into

$$\frac{dx}{\sqrt{-(ax^2+2bx+c)(a''x^2+2b''x+c'')}} = \frac{dy}{\sqrt{(by^2+b'')^2-(ay^2+a'')(cy^2+c'')}}. \tag{34}$$

We have only to factor the polynomial of degree four in (26) into two second degree factors, and then equate the coefficients of the left hand side of (34) to the corresponding coefficients in (26). In this way we obtain the theorem due to Lagrange* that after the differential (26) has been brought to the form

$$(35)\qquad \frac{dx}{\sqrt{-(ax^2+2bx+c)(a''x^2+2b''x+c'')}},$$

the transformation

$$(36)\qquad y^2 = -\frac{a''x^2+2b''x+c''}{ax^2+2bx+c}$$

will carry it over into another differential

$$(37)\qquad \frac{dy}{\sqrt{\alpha+\beta y^2+\gamma y^4}}$$

which contains only even powers of y. The constants α, β and γ are determined by

$$(38)\qquad \alpha = b''^2 - a''c'',\quad \beta = 2bb' - ac'' - ca'',\quad \gamma = b^2 - ac.$$

This reduced form (36) of the differential (26) contains only *two* essential constants. Putting $\sqrt{\alpha y}$ instead of y and setting

$$\frac{\beta}{\alpha} = -(p^2+q^2),\quad \frac{\gamma}{\alpha} = p^2q^2,$$

(36) is carried over into

$$(39)\qquad \frac{dy}{\sqrt{(1-p^2y^2)(1-q^2y^2)}}.$$

The differential (26), *which contains five independent constants, can be transformed by means of a substitution of the form* (36) *into another differential* (39)

* Jacobi erroneously credits Legendre with this transformation (Math. Werke, vol. 3, p. 85). It is to be found in Lagrange's paper "Sur une nouvelle méthode de calcul intégral etc.", published in the Proceedings of the Academy of Sciences of Turin (1784-85). Cf. Oeuvres de Lagrange, vol. 2, p. 256.

which contains only two constants. This important theorem was discovered by Euler but he performs the transformation of (26) into (39) by means of an equation which is of the *first* degree in both y and x.

Thus we have seen how Euler carried the differential (26) over into another differential (27) of the same form and with the same constants by means of a transformation of the general form (29). From this transformation arose *Euler's addition theorem.* We have seen further how Lagrange and Euler, using other transformations also of the general form (29), reduced the differential (26) to the normal form (39).

However, this does not exhaust the importance of the transformation (29). To this general form belongs also the peculiar *Gaussian transformation* by means of which we can transform the normal form (39) into another differential of the *same* form but with *different* constants.

If we put

$$x = \frac{2y}{1+q^2y^2} \tag{40}$$

$$\begin{aligned} p_1 &= p+\sqrt{p^2-q^2}, & p &= \tfrac{1}{2}(p_1+q_1), \\ q_1 &= p-\sqrt{p^2-q^2}, & q &= \sqrt{p_1 q_1}, \end{aligned} \tag{41}$$

we obtain

$$dx = \frac{2(1-q^2y^2)}{(1+q^2y^2)^2}\,dy.$$

$$\begin{aligned} 1-p^2x^2 &= \frac{(1-p_1^2y^2)(1-q_1^2y^2)}{(1+q^2y^2)^2}, \\ 1-q^2x^2 &= \left(\frac{1-q^2y^2}{1+q^2y^2}\right)^2. \end{aligned} \tag{42}$$

and consequently

$$\begin{aligned} \frac{dx}{\sqrt{(1-p^2x^2)(1-q^2x^2)}} &= \frac{2\,dy}{\sqrt{(1-p_1^2y^2)(1-q_1^2y^2)}} \\ &= \frac{d(2y)}{\sqrt{[1-(\frac{1}{2}p_1)^2(2y)^2]\,[1-(\frac{1}{2}q_1)^2(2y)^2]}}. \end{aligned} \tag{43}$$

If, for a differential of the form (39), we write the multiplication theorem according to Euler with $n = 2$, we find that the middle term in equation (43) is transformed into

$$\frac{dy_1}{\sqrt{(1-p_1^2y_1^2)(1-q_1^2y_1^2)}} \tag{44}$$

by the transformation

$$y_1 = \frac{2y}{1-p_1^2 q_1^2 y^4} \sqrt{(1-p_1^2 y^2)(1-q_1^2 y^2)}. \tag{45}$$

The last equation may also be written

$$y_1 = \frac{2y}{1+q^2 y^2} \sqrt{\frac{(1-p_1^2 y^2)(1-q_1^2 y^2)}{(1-q^2 y^2)^2}}.$$

Thus, in view of (40) and (42), the differential (39) is transformed into the differential (44) by means of the transformation

$$y_1 = x \sqrt{\frac{1-p^2 x^2}{1-q^2 x^2}}. \tag{46}$$

Hence, we obtain the following two theorems, namely,

A. a) The differential

$$\frac{dx}{\sqrt{(1-p^2 x^2)(1-q^2 x^2)}}$$

is carried over into a new differential of the same form

$$\frac{dy}{\sqrt{[1-(\frac{1}{2} p_1)^2 y^2][1-(\frac{1}{2} q_1)^2 y^2]}}$$

by the transformation

$$x = \frac{y}{1+\frac{1}{4} q^2 y^2}.$$

x and y vanish simultaneously. Instead of the old constants p and q, new constants, $\frac{1}{2} p_1$ and $\frac{1}{2} q_1$, appear which are given by the relations

$$\tfrac{1}{2} p_1 = \tfrac{1}{2}(p+\sqrt{p^2-q^2}), \qquad \tfrac{1}{2} q_1 = \tfrac{1}{2}(p-\sqrt{p^2-q^2}).$$

If we replace y, $\frac{1}{2} p_1$ and $\frac{1}{2} q_1$ by x, p and q respectively and p and q by $2 p_1$ and $2 q_1$, we obtain

A. b) The differential

$$\frac{dx}{\sqrt{(1-p^2x^2)(1-q^2x^2)}}$$

is carried over into a new differential of the same form

$$\frac{dy}{\sqrt{(1-4p_1^2y^2)(1-4q_1^2y^2)}}$$

by means of the transformation

$$y = \frac{x}{1+pqx^2}.$$

x and y vanish simultaneously. Instead of the old constants p and q, new constants, p_1 and q_1, appear which are given by the relations

$$p_1 = \tfrac{1}{2}(p+q), \qquad q_1 = \sqrt{pq}.$$

Similarly formula (46) will give us two transformation theorems. The first of these is obtained by replacing y_1, p_1 and q_1, by x, p and q respectively.

B. a) The differential

$$\frac{dx}{\sqrt{(1-p^2x^2)(1-q^2x^2)}}$$

is carried over into a differential of the same form

$$\frac{dy}{\sqrt{(1-p_1^2y^2)(1-q_1^2y^2)}}$$

by means of the transformation

$$x = y\sqrt{\frac{1-[\frac{1}{2}(p+q)]^2y^2}{1-pqy^2}}.$$

x and y vanish simultaneously. Instead of the old constants p and q new constants p_1 and q_1 appear which are given by the relations

$$p_1 = \tfrac{1}{2}(p+q), \qquad q_1 = \sqrt{pq}.$$

B. b) The differential

$$\frac{dx}{\sqrt{(1-p^2x^2)(1-q^2x^2)}}$$

is carried over into a new differential of the same form

$$\frac{dy}{\sqrt{(1-p_1^2y^2)(1-q_1^2y^2)}}$$

by means of the transformation

$$y = x\sqrt{\frac{1-p^2x^2}{1-q^2x^2}}.$$

x and y vanish simultaneously. Instead of the old constants p and q, new constants, p_1 and q_1, appear which are given by the relations

$$p_1 = p+\sqrt{p^2-q^2}, \qquad q_1 = p-\sqrt{p^2-q^2}.$$

We have seen how the first of these two transformation theorems, theorem A, is contained as a particular case of the significant transformation (29). We have seen further how the second, theorem B, can be obtained from the former by a simple application of the multiplication theorem which also is deduced from the same transformation. The discovery contained in these two theorems was a new conquest of fundamental importance for the theory with which we are concerned. Once a normal form of the differential

$$\frac{dx}{\sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}}$$

has been found containing the minimum number of independent constants, it is possible to transform this normal form further without increasing the number of the constants.

The transformation theorem B was discovered in 1771 by Landen* who does not seem, however, to have fully understood the value of his discovery.

* Philosophical Transactions, (1771). Landen's investigations are further developed in "On the Investigation of a General Theorem for finding the Length of any Arc of any Conic Hyperbola, by Means of Two Elliptic Arcs, with some other new and useful Theorems deduced therefrom." Philosophical Transactions, vol. 65, (1775) p. 283, and in "Of the Ellipsis and Hyperbola," Mathematical Memoirs by John Landen, F. R. S., London, 1780, p. 23.

Lagrange* was the first to correctly appreciate the importance of Landen's work. He tried to obtain a transformation of

$$\int \frac{dx}{\sqrt{(1-p^2x^2)(1-q^2x^2)}}$$

into another differential of the same form but with different constants, his purpose being to obtain means for an approximate calculation of the integral

$$\int \frac{dx}{\sqrt{(1-p^2x^2)(1-q^2x^2)}}$$

by repeated transformations in order to make the integral approach an arc sine indefinitely. With this purpose in view he proposed the transformation theorem B, and it is probable that he had discovered this theorem independently without knowledge of Landen's previous work. These investigations by Lagrange are contained in the Transactions of the Academy of Sciences of Turin for the years 1784–85.† The transformation theorem A, though simpler, was not published until 1818 by Gauss who used it for solving a mechanical problem.‡ *Euler's addition theorem and the transformation theorem of Landen and Lagrange* were the two fundamental ideas of which the theory of elliptic functions was in possession when this theory was brought up for renewed consideration by Legendre§ in 1786. During four decades he was the only man who added investigations concerning these integrals to mathematical literature. His broad investigations are collected in two big publications: *Exercises de Calcul Intégral* and *Traité des Fonctions Elliptiques.*

Granting the merits of Legendre's‖ work, and the great importance one has

* Cf. Enneper: Elliptische Funktionen, Theorie und Geschichte, Halle (1876), p. 307. This textbook is full of historical information.

† "Sur une nouvelle méthode de Calcul Intégral pour les différentielles affectées d'un radical carré sous lequel la variable ne passe pas le quatrième degré", Mémoires de l'Académie royale des Sciences de Turin, vol. 2 (1784–85); Oeuvres de Lagrange, vol. 2, p. 253.

‡ "Determinatio Attractionis quam in punctum quodvis positionis datae exerceret planeta si ejus massa per totam orbitam ratione temporis quo singulae partes discribuntur uniformiter esset dispertita", Commentationes soc. reg. scient. Gottingensis, vol. 4 (1818), p. 21-48; also Werke, vol. 3, p. 352, et seq.

§ "Mémoire sur les Intégrations par Arcs d'Ellipse", Histoire de l'Académie Royale des Sciences (1786), pp. 616-643; "Second Mémoire sur les Intégrations par Arcs d'Ellipse", Hist. de l'Acad. des Sciences (1786), pp. 644-683.

‖ In addition to the memoirs of 1786, Legendre has published the following investigations on elliptic functions, namely,

(1) "Mémoire sur les transcendentes elliptiques, lu à la ci-devant Académie des Sciences avril 1792", Paris, L'an deuxième de la République.

(2) "Exercises de Calcul Intégral sur divers ordres de transcendents et sur les Quadratures",

to attach to it, it must be admitted that he did not add any new mathematical ideas of fundamental importance to the discoveries of Euler and Lagrange. He has drawn many conclusions, previously unstated, out of the fundamental concepts which we have expounded above, and on that foundation he has erected a mathematical theory of great extent. Still, the real intrinsic nature of the functions he was engaged in studying escaped his notice altogether.

One of the chief merits of his investigations is that he showed that all elliptic integrals could be reduced to three fixed canonical forms which in trigonometric form can be written as follows

$$F(\varphi) = \int_0^\varphi \frac{d\varphi}{\sqrt{1-k^2\sin^2\varphi}},$$

$$E(\varphi) = \int_0^\varphi \sqrt{1-k^2\sin^2\varphi}\, d\varphi,$$

$$\Pi(\varphi) = \int_0^\varphi \frac{d\varphi}{(1+n\sin^2\varphi)\sqrt{1-k^2\sin^2\varphi}}.$$

But the most profound idea which is due to Legendre in this field of mathematics is that he raised the question whether it was possible to propose a more general transformation problem than that of Landen, Lagrange, and Gauss. The transformation of Gauss consists in replacing the variable x by a new variable y which is a certain rational function of x of the *second* degree; Legendre tried to generate another similar transformation by letting y be a rational function of x of the *third* degree, and actually succeeded in finding

Paris (1811). A second volume of this book appeared in 1817 and the year before, 1816, a third volume. In the introduction of the second volume Legendre mentions that he has published a supplement to the first volume in 1811.

(3) "Traité des fonctions Elliptiques et des Intégrales Eulériennes. Tome premier, contenant la théorie des Fonctions elliptiques et son application à different problèmes de Géométrie et de Mécanique", Paris (1825). "Tome second, contenant les méthodes pour construire les Tables elliptiques, le Receuil de ces Tables, le Traité des integrales Eulériennes et un Appendice", Paris (1826). Also on account of the publications of Abel and Jacobi: "Premier Supplément et Deuxième Supplément" (the latter dated March 15, 1829), and "Troisième Supplément", (dated March 4, 1832).

(4) "Note sur les nouvelles propriétés des Fonctions elliptiques découvertes par M. Jacobi", dated February 11, 1828 and published in Astronom. Nachrichten, vol. 6, no. 130, pp. 201-208; this number appeared in Altona in February 1828.

such a transformation.* It seems, however, as if he was prevented from considering the general problem of finding all possible transformations in which y is a rational function of x of arbitrary degree by relying on certain peculiar geometrical intuitions which are really foreign to the nature of the subject.† But even if he had set himself this question, and even if he had managed to prove the possibility of solving it, the problem of actually representing this solution was far from the stand-point occupied by Legendre. This required the introduction of certain essentially new fundamental concepts, the discovery of which was reserved for Abel and Jacobi.

These new discoveries were presented to the scientific world through the simultaneous publications of Abel and Jacobi in September 1827, a date that will always mark an epoch in the history of mathematics.‡

Jacobi had penetrated deeply into the general transformation problem which Legendre had touched upon without recognizing its full significance; and his endeavor to solve this problem led him to the new theory of *elliptic functions*.

The profound algebraic investigations of Abel had led him to seek the answer to another question. Gauss had indicated as early as 1801 in his *Disquisitiones Arithmeticae*§ that the results published concerning the division of the circle could be extended to hold for the *lemniscate*. The differential of the arc of a lemniscate belongs to the general form (26). The multiplication theorem of Euler showed how to transform the n^{th} multiple of a differential in x, of the form (26), into another differential in y, (27), of the same form and with the same constants. This transformation was effected by equating y to a certain rational function of x of degree n^2, whereas in the corresponding problem for the inverse trigonometric functions this rational function was of

* Cf. Traité des fonctions elliptiques, chapt. 31. It appears from the correspondence between Jacobi and Legendre, and also from a foot-note in Jacobi's first note in Schumacher's Journal, that Jacobi was not cognizant of Legendre's transformation at the time of the publication of the note mentioned. As a matter of fact, this transformation was published for the first time in the first volume of Traité d. fonct. ell. It is true that this volume was printed in 1825, but the publication as a whole was not available at the booksellers' until January 1827. Cf. "Correspondance mathématique entre Legendre et Jacobi", Journal für Mathematik, vol. 80, (1875), p. 215.

† See the page quoted in vol. 80 of Journal f. Math.

‡ Abel, "Recherches sur les fonctions elliptiques", Journal für Mathematik, vol. 2, no. 2, p. 101. According to Borchardt — ibid., vol. 80, p. 206 — this number appeared in Berlin in September 1827.

Jacobi, "Extraits de deux lettres de M. Jacobi de l'Université de Königsberg à l'éditeur", Astronomische Nachrichten, vol. 6, no. 123, pp. 33-38. This number appeared in Altona in September 1827.

§ [Loc. cit. Maser's German edition, Berlin (1889), p. 397. Translator's note.] Werke vol. 1, p. 412.

degree n. Thus the division problem for the *lemniscate* depends upon the solution of an algebraic equation of degree n^2 while the corresponding division problem for the *circle* depends upon an algebraic equation of degree n. In the latter case the meaning of the n different roots was easily interpreted; in the former case, on the other hand, the meaning of the roots was quite puzzling as the greater part of them were complex numbers. It was probably the endeavor to perform this extension of the division problem, indicated by Gauss, and at the same time to explain the difficulties mentioned above, which led Abel to the discovery of *the elliptic functions*.*

The reference in *Disquisitiones arithmeticae* mentioned above shows that Gauss as early as 1801 had some knowledge of quite a few of the theorems which were later found by Abel and Jacobi. The correctness of Gauss' declarations by word of mouth in letters† that he had been in possession of the elliptic functions and their chief properties from the end of the eighteenth century, has been fully proved by Schering's publications of Gauss' posthumous papers.‡

This ends our historical survey of the investigations which prepared the way for the discovery of the theory of the elliptic functions. It follows from Euler's addition theorem that the simplest of the elliptic differentials is given by

$$du = \frac{dx}{\sqrt{Ax^4 + 4Bx^3 + 6Cx^2 + 4B'x + A'}};$$

hence the efforts to integrate this differential equation became the starting point of the new theory. Thus when we proceed to give a presentation of the main roads along which one can arrive at *the elliptic functions*, it is natural for us to begin with a description of the roads which start from the study of this differential equation. This is not only historically the oldest procedure, but it leads also to the goal with a minimum use of assumptions from function theory and thus gives the closest connection with elementary mathematics.

* Œuvres complètes de N. H. Abel rédigées par Holmboe; Extraits de quelques lettres de l'auteur à l'éditeur, vol. 2, p. 271. [Same volume, p. 262, in the new edition by Sylow and Lie. Translator's note.]

Lejeune-Dirichlet, "Gedächtnisrede auf Carl Gustav Jacob Jacobi", Journal f. Math., vol. 52, (1856), p. 199.

† Consult especially statements in letters to Schumacher: "Briefwechsel zwischen C. F. Gauss und H. C. Schumacher," herausgegeben von A. F. Peters. [Also consult Gauss' letter to Bessel of March 30, 1828 in "Briefwechsel zwischen Gauss und Bessel", Leipzig, (1880), p. 477, reproduced in Gauss' Werke, vol. 10:1, p. 248, T. n.]

‡ Gauss' Werke, vol. 3, p. 231 et seq. [Further cf. Werke, vol. 10:1 Analysis, especially pp. 145–325, also many entries in the diary, reproduced on pp. 485–574. The first entry on lemniscatic functions is dated January 8, 1797. Translator's note.]

§ 1.

The reduction of the differential equation

$$(1) \qquad \frac{dx}{\sqrt{R(x)}} = du,$$

$$(2) \qquad R(x) = Ax^4 + 4Bx^3 + 6Cx^2 + 4B'x + A',$$

to the normal form of Weierstrass and the deduction of the addition theorem.

We make the following assumptions concerning the constants entering in $R(x)$. The constant A may be zero, but B shall not vanish simultaneously with A. Further the roots of $R(x) = 0$ shall all be unequal; otherwise the differential equation (1) would reduce to an equation whose solution would be a trigonometric or algebraic function.

Turning to the problem of finding the functional relation between the variables x and u, which satisfy equation (1), the first problem we have to solve is to reduce the equation to the simplest possible form.

We shall first find out what simplifications can be effected by the introduction of a new variable x' instead of x such that a bilinear relation holds between x and x'. The general form of such a relationship is given by

$$(3) \qquad (cx' + d)x + ax' + b = 0,$$

or

$$x = \frac{\alpha x' + \beta'}{x' + \delta}.$$

Then we have

$$dx = \frac{\alpha\delta - \beta}{(x' + \delta)^2}\, dx'.$$

We put

$$(4) \qquad R_1(x') = \frac{(x' + \delta)^4}{(\alpha\beta - \delta)^2}\, R\left(\frac{\alpha x' + \beta}{x' + \delta}\right).$$

Consequently $R_1(x')$ is an integral rational function of x' of degree four.

If we determine the sign of $\sqrt{R_1(x')}$ in such a way that the relation

$$(5) \qquad \sqrt{R_1(x')} = -\frac{(x' + \delta)^2}{\alpha\beta - \delta}\sqrt{R(x)}$$

holds, we obtain

$$\frac{dx}{\sqrt{R(x)}} + \frac{dx'}{\sqrt{R_1(x')}} = 0. \tag{6}$$

Thus the transformation (3) carries the differential equation (1) into a new differential equation

$$-\frac{dx'}{\sqrt{R_1(x')}} = du. \tag{7}$$

We shall now try to simplify this transformed equation by an appropriate choice of the constants α, β and δ. The equation (4) can be written

$$R_1(x') = \frac{(x'+\delta)^4}{(\alpha\beta-\delta)^2} R\left(\alpha + \frac{\beta-\alpha\delta}{x'+\delta}\right)$$

$$= \frac{R(\alpha)}{(\beta-\alpha\delta)^2}(x'+\delta)^4 + \frac{R'(\alpha)}{\beta-\alpha\delta}(x'+\delta)^3 + \frac{1}{2!}R''(\alpha)(x'+\delta)^2 \tag{8}$$

$$+ \frac{1}{3!}R'''(\alpha)(\beta-\alpha\delta)(x'+\delta) + \frac{1}{4!}R^{(IV)}(\alpha)(\beta-\alpha\delta)^2.$$

This formula shows immediately:

(1) The coefficient of the fourth power of x' in $R_1(x')$ can be made to vanish by equating α to one of the roots of $R(x) = 0$.

(2) It is not possible to make the coefficient of the third power of x' vanish simultaneously with the coefficient of the fourth power.

(3) α being determined through the equation $R(\alpha) = 0$, we can make the coefficient of the second power of x' vanish by introducing a certain linear and homogeneous relation between β and δ.

(4) We can give the coefficient of the third power of x' a definite numerical value by introducing still another linear relation between β and δ.

Thus, by means of the linear transformation (3), we can transform the fundamental differential equation (1) into a new differential equation (7) which contains only *two* constants instead of the previous *five*. The coefficients of x'^4 and x'^2 are zero and the coefficient of x'^3 has a definite numerical value which is independent of the *five* constants in $R(x)$.

One can look at the equation (3) from another point of view, namely as an integral of the differential equation (6). This integral is a *particular* one,

since to every given value of one variable corresponds a perfectly definite value of the other. Then the question arises: Is it possible to generalize this *particular* integral in such a way that it becomes a complete integral? Or in other words, is it possible to find a relation between the variables x and x' which satisfies the equation (6), and, moreover, such that the two variables can be given arbitrary initial values? A *complete* integral of our fundamental equation

$$\frac{dx}{\sqrt{R(x)}} = du$$

must permit the designation of arbitrary initial values for the variables x and u. If it is possible to find a relation between x and x' of the type mentioned above, much would be gained. On the one hand the differential equation

$$\frac{dx}{\sqrt{R(x)}} = du,$$

which contains *five* independent constants, would be reduced to a new differential equation

$$-\frac{dx'}{\sqrt{R_1(x')}} = du,$$

which contains only *two* constants. On the other hand, after the initial values of u and x' have been fixed once for all, an *arbitrary* initial value of x could always be made to correspond to the same initial value of x'. Thus the investigation of the *complete* integral of our differential equation (1), would be reduced to an investigation of a certain *particular* integral of the reduced differential equation (7). In the former case we have to do with a functional relation between two variables x and u, involving *six* independent constants; in the latter case our problem would be restricted to the consideration of a functional relation, involving only *two* constants. Thus the simplification to be gained by the discovery of the *complete* integral of the differential equation (6) is considerable.

We know that this equation admits of a *particular* integral which is an entire rational function of the *first* degree in the variables x and x'. Hence it is reasonable to look for the *complete* integral among the entire rational functions of the *second* degree in the variables x and x'.

The general equation of the second degree between two variables can be written

$$(9) \qquad Lx'^2 + Mx' + N = L_1 x^2 + M_1 x + N_1 = 0$$

where L, M and N are algebraic functions in x of the second degree and L_1, M_1 and N_1 are integral algebraic functions in x', also of the second degree.

From (9) we obtain

$$(10)\qquad \begin{aligned} 2Lx'+M &= \sqrt{M^2-4LN}\,, \\ 2L_1x+M_1 &= \sqrt{M_1^2-4L_1N_1}\,, \end{aligned}$$

and consequently

$$(11)\qquad \frac{dx}{\sqrt{M^2-4LN}} = -\frac{dx'}{\sqrt{M_1^2-4L_1N_1}}\,.$$

The functions under the radicals are at most of degree four. Thus if we want to make the differential equation (11) identical with the differential equation (6), we must be able to determine the *nine* constants in (9) and a *tenth* constant k^2 such that the following equalities hold, namely

$$(12)\qquad M^2-4LN = k^2R(x),$$

$$(13)\qquad M_1^2-4L_1N_1 = k^2R_1(x').$$

If these two equalities are to hold, it is necessary first that the different powers of x have the same coefficients on both sides of equation (12). This implies *five* algebraic equations between our *ten* constants. Further, since in $R_1(x')$ the coefficients of the fourth and the second powers of x' must be zero and the coefficient of x'^3 has to have a certain numerical value, we have *three* more algebraic equations between the *ten* constants. Thus we have altogether *eight* algebraic equations between our *ten* constants. A glance shows that all the equations are homogeneous and of the second degree in the ten constants. Hence we can give one of the constants a suitable numerical value, and are still left with *eight* algebraic equations between *nine* otherwise independent constants. Thus *one* constant remains arbitrary. If this constant enters in our algebraic conditional equations in such a manner that we can assign *arbitrary* initial values to the associated variables x and x', then the equation (9) is the *complete* integral of (6).

We write*

$$(14)\qquad L = \lambda x^2+\mu x+\nu.$$

* The elegant calculations which we are going to carry out, are essentially the same as those which Weierstrass used to give in his lectures. [Cf. Weierstrass' Mathematische Werke, vol. 5, Chapt. 1. T. n.]

One of our eight algebraic equations has to express that the coefficient of the fourth power of x' in $R_1(x')$ is zero. In view of (9) and (13) this equation is given by

(15) $$\mu^2 - 4\lambda\mu = 0.$$

Therefore, it follows that

$$L = \lambda\left(x + \frac{\mu}{2\lambda}\right)^2.$$

On account of the homogeneity of the eight algebraic equations we can give an arbitrary value to one of the ten constants. Let us set

(16) $$\lambda = 1,$$

obtaining

$$L = \left(x + \frac{\mu}{2}\right)^2.$$

Instead of μ we introduce a new constant x_0 by setting

(17) $$-\frac{\mu}{2} = x_0.$$

Consequently

(18) $$L = (x - x_0)^2.$$

Now we proceed to deduce the *five* conditions arising out of (12). We have

$$N = \frac{M^2 - k^2 R(x)}{4L}.$$

Since N is an integral function in x, $M^2 - k^2 R(x)$ must be exactly divisible by L or $(x - x_0)^2$. Thus the coefficients of the first and the zero power of $x - x_0$ in $M^2 - k^2 R(x)$ must be zero. This gives *two* of the *five* equations arising from (12). We can write $R(x)$ in the form

(19) $$R(x) = \sum_{n=0}^{4} r_n (x - x_0)^n.$$

where r_n $(n = 0, 1, \ldots. 4)$ are functions of x_0. Further we can write M in the form

$$M = \sum_{n=0}^{2} m_n (x - x_0)^n. \tag{20}$$

Thus the two equations mentioned become

$$m_0^2 - k^2 r_0 = 0, \qquad 2 m_0 m_1 - k^2 r_1 = 0. \tag{21}$$

Instead of m_0 and m_2 we introduce the quantities g and h, determined by the equations

$$m_0 = -\frac{r_0}{g}, \qquad m_0 = -\frac{h}{g}. \tag{22}$$

thus obtaining

$$k^2 = \frac{r_0}{g^2}, \qquad m_1 = -\frac{r_1}{2g}, \tag{23}$$

and consequently

$$M = -\frac{1}{g}\left[r_0 + \frac{r_1}{2}(x - x_0) + h(x - x_0)^2\right], \tag{24}$$

and, dividing L into $M^2 - K^2 R(x)$,

$$\begin{aligned} N = \frac{1}{4g^2} [(\tfrac{1}{4} r_1^2 + 2 r_0 h - r_0 r_2) + (h r_1 - r_0 r_3)(x - x_0) \\ + (h^2 - r_0 r_4)(x - x_0)^2]. \end{aligned} \tag{25}$$

Thus we have determined the three coefficients of N in terms of previously introduced quantities, and the *five* equations which arise out of (12) are contained in (23) and (25).

It is easy to determine L_1, M_1, N_1 by means of the expressions for L, M and N, given in (18), (24) and (25). However, instead of L_1, M_1, N_1 we introduce three new functions L', M', N', determined by

$$Lx'^2 + Mx' + N = L_1 x^2 + M_1 x + N_1 = L'(x - x_0)^2 + M'(x - x_0) + N'. \tag{26}$$

This enables us to write

$$L' = x'^2 - \frac{h}{g}x' + \frac{h^2 - r_0 r_4}{4g^2} = \left(x' - \frac{h}{2g}\right)^2 - \frac{r_0 r_4}{4g^2},$$

(27)
$$M' = -\frac{r_1}{2g}x' + \frac{hr_1 - r_0 r_3}{4g^2} = -\frac{r_1}{2g}\left(x' - \frac{h}{2g}\right) - \frac{r_0 r_3}{4g^2},$$

$$N' = -\frac{r_0}{g}x' + \frac{1}{4g^2}(\tfrac{1}{4}r_1^2 + 2r_0 h - r_0 r_2)$$

$$= -\frac{r_0}{g}\left(x' - \frac{h}{2g}\right) + \frac{r_0 r_2 - \frac{1}{4}r_1^2}{4g^2}.$$

Observing that

(28)
$$M_1^2 - 4L_1 N_1 = M'^2 - 4L'N',$$

we see from (13) that

(29)
$$R(x) = 4g\left(x' - \frac{h}{2g}\right)^3 + r_2\left(x' - \frac{h}{2g}\right)^2$$
$$+ \frac{r_1 r_3 - 4r_0 r_4}{4g}\left(x' - \frac{h}{2g}\right) + \frac{r_0 r_3^2 + r_4 r_1^2 - 4r_0 r_2 r_4}{16g^2}.$$

So far we have introduced *six* of our *eight* algebraic equations. The remaining *two* express, one that the coefficient of x'^3 is a certain number, the other that the coefficient of x'^2 is zero. It appears from the form of the coefficient of the third power of x' in (29) that a simple way of choosing this number is to equate it to 4. Then we have

(30)
$$g = 1.$$

The remaining equation whereby the coefficient of the second power reduces to zero is $-6h + r_2 = 0$, or

(31)
$$h = \frac{r_2}{6}.$$

Consequently, we have

(32)
$$R_1(x') = 4x'^3 - g_2 x' - g_3,$$

where

$$
(33)\qquad
\begin{aligned}
g_2 &= r_0\, r_4 - \tfrac{1}{4}\, r_1\, r^3 + \tfrac{1}{12}\, r_2^2,\\
g_3 &= \tfrac{1}{6}\, r_0\, r_2\, r_4 + \tfrac{1}{48}\, r_1\, r_2\, r_3 - \tfrac{1}{16}\, r_0\, r_3^2 - \tfrac{1}{16}\, r_4\, r_1^2 - \tfrac{1}{216}\, r_2^3.
\end{aligned}
$$

The equations (23) and (25) represent the *five* equations which we obtain by equating the coefficients on both sides of the equation (12). By (16) we have made an arbitrary determination of one of the ten constants which were connected by eight homogeneous equations at the start of our investigation. By means of the equations (15), (30) and (31) we have made the coefficients of the fourth and the second powers of x' in $R_1(x')$ equal to zero and given the coefficient of the third power the value 4.

By means of these equations the *nine* essential coefficients of the problem are expressed in terms of the coefficients of $R(x)$ and an arbitrary quantity x_0. Thus the two constants which enter in $R_1(x')$ are expressed in terms of the coefficients of $R(x)$ and *in terms of* x_0. However, the dependence upon x_0 is only apparent; differentiating (19) with respect to x_0 and observing that this derivative is identically zero, no matter what value x has, we obtain

$$
(34)\qquad
\begin{aligned}
&r_0 = R(x_0), \qquad r_1 = \frac{dr_0}{dx_0}, \qquad 2\,r_2 = \frac{dr_1}{dx_0},\\
&3\,r_3 = \frac{dr_2}{dx_0}, \qquad 4\,r_4 = \frac{dr_3}{dx_0}, \qquad \frac{dr_4}{dx_0} = 0,
\end{aligned}
$$

and consequently

$$
(35)\qquad \frac{dg_2}{dx_0} = 0, \qquad \frac{dg_3}{dx_0} = 0.
$$

Thus we can express g_2 and g_3 in terms of the coefficients of $R(x)$ by simply putting $x_0 = 0$ in (33). We then obtain

$$
(36)\qquad
\begin{aligned}
g_2 &= AA' + 3\,C^2 - 4\,BB'.\\
g_3 &= ACA' + 2\,BCB' - AB'^2 - A'B^2 - C^3.
\end{aligned}
$$

Instead of x' and $R_1(x')$ we introduce the notation of Weierstrass, namely

$$
s = x', \qquad S = R_1(x').
$$

obtaining

$$S = 4s^3 - g_2 s - g_3. \tag{37}$$

In view of (10) and (34), a simple calculation gives

$$\begin{aligned} s &= \frac{\sqrt{R(x_0)}\sqrt{R(x)} + R(x_0) + \frac{1}{2}R'(x_0)(x-x_0)}{2(x-x_0)^2} + \tfrac{1}{24}R''(x_0) \\ &= \frac{1}{4}\left\{\frac{\sqrt{R(x)} + \sqrt{R(x_0)}}{x-x_0}\right\}^2 - \tfrac{1}{4}A(x+x_0)^2 - B(x+x_0) - C. \end{aligned} \tag{38}$$

This equation satisfies the differential equation

$$\frac{dx}{\sqrt{Ax^4 + 4Bx^3 + 6Cx^2 + 4B'x + A'}} = \frac{ds}{\sqrt{4s^3 - g_2 s - g_3}}, \tag{39}$$

if only the meaning of the radical on the right hand side is correctly determined. It further contains an arbitrary constant x_0 which can be determined in such a way that arbitrarily chosen initial values can be assigned for s and x.

Thus we have succeeded in solving our problem and have obtained the *complete* integral of equation (39) in the expression (38). A glance at (38) shows that, if we let x_0 be the arbitrary initial value of x, the variable s will always receive the same initial value, namely $s_0 = \infty$.

Differentiating the first formula in (38) with respect to x and using formula (39), we obtain further

$$\begin{aligned} \sqrt{S} &= -\left\{\frac{R'(x)}{4(x-x_0)^2} - \frac{R(x)}{(x-x_0)^3}\right\}\sqrt{R(x_0)} \\ &\quad + \left\{\frac{R'(x_0)}{4(x-x_0)^2} + \frac{R(x_0)}{(x-x_0)^3}\right\}\sqrt{R(x)}, \end{aligned} \tag{40}$$

which shows, among other things, which of the two values of $\sqrt{S}$ corresponds to preassigned values of $\sqrt{R(x)}$ and $\sqrt{R(x_0)}$.

By means of the equations (39) *and* (40), *s and $\sqrt{S}$ are expressed rationally in terms of x and $\sqrt{R(x)}$. From* (9) *and* (10) *it follows that we can also express x and $\sqrt{R(x)}$ rationally in terms of s and $\sqrt{S}$.* We leave out the actual development of these formulas since they are not of essential importance to the solution of our main problem.

Thus we obtain the following theorem:

The differential equation

$$\frac{dx}{\sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}} = du,$$

containing five independent constants A, B, C, B', A', *in which the variable* x *is given the arbitrary initial value* x_0, *is transformed by means of equation* (38) *into a new differential equation*

$$-\frac{ds}{\sqrt{4s^3-g_2s-g_3}} = du,$$

containing only two constants g_2 *and* g_3. *The constants are rational, integral functions,* (36), *of the constants mentioned above and, to the arbitrary initial value* x_0 *of* x, *always corresponds the value* ∞ *of* s. *The equation* (40) *shows which of the two values of* $\sqrt{S}$ *corresponds to preassigned values of* $\sqrt{R(x)}$ *and* $\sqrt{R(x_0)}$.

We do not restrict the generality of our conclusions by always assigning *zero* as the initial value of the variable u.

This theorem can also be expressed as follows:

If it is possible to find a functional relation between the variables s *and* u, *satisfying the differential equation*

$$-\frac{ds}{\sqrt{4s^3-g_2s-g_3}} = du,$$

which is such that s *becomes infinite when* u *becomes zero, then we have at the same time obtained a functional relation between the variables* x *and* u *which satisfies the differential equation*

$$\frac{dx}{\sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}} = du$$

and is such that the variable x *takes on the arbitrary value* x_0 *for* $u=0$. x *is a certain rational function of* s, $\sqrt{S}$, x_0 *and* $\sqrt{R(x_0)}$ *which for an infinite value of* s *takes on the value* x_0. *Thus we need only to express* s *in terms of* u *in order to obtain the functional relation between the variables* x *and* u. *From the equation* (40) *we find which of the two values of* $\sqrt{S}$ *to use.*

It is now an easy matter to obtain Euler's addition theorem directly.* Euler set himself the task to find the complete integral of the differential equation

$$\frac{dx}{\sqrt{R(x)}} = \frac{dx'}{\sqrt{R(x')}}, \tag{41}$$

which can be replaced by the two simultaneous equations

$$\begin{aligned} \frac{dx}{\sqrt{R(x)}} &= du = -\frac{ds}{\sqrt{S}}, \\ \frac{dx'}{\sqrt{R(x')}} &= du = -\frac{ds}{\sqrt{S}}. \end{aligned} \tag{42}$$

Now, if we choose arbitrarily the initial values x_0 and x_0' and the quantity x and determine the corresponding values of the radicals, we find from the relations (38) and (40) that s and S are rational of functions x, x_0, $\sqrt{R(x)}$ and $\sqrt{R(x_0)}$. Likewise we find that x' is a rational function of s and $\sqrt{S}$ and of x_0' and $\sqrt{R(x_0')}$. Consequently x' becomes a rational function of x_0', x_0, x and $\sqrt{R(x_0')}$, $\sqrt{R(x_0)}$, $\sqrt{R(x)}$; this function satisfies the equation (41) and has, moreover, the property that x takes on the arbitrary value x_0, when x' takes on the arbitrary value x_0'.

Thus we have obtained the following theorem:

The differential equation

$$\frac{dx}{\sqrt{R(x)}} = \frac{dx'}{\sqrt{R(x')}},$$

or the simultaneous differential equations

$$\frac{dx}{\sqrt{R(x)}} = du, \qquad \frac{dx'}{\sqrt{R(x')}} = du,$$

are satisfied by a certain definite algebraic function of x and x' such that x' is rationally expressible in terms of x, and the corresponding value of $\sqrt{R(x)}$, the arbitrary constants x_0, x_0' and the corresponding $\sqrt{R(x_0)}$ and $\sqrt{R(x_0')}$. x' takes on the arbitrarily preassigned value x_0' when x takes on the arbitrary

* This remark is due to Weierstrass.

value x_0. Likewise, $\sqrt{R(x')}$ is a rational function of x, $\sqrt{R(x)}$, x_0, $\sqrt{R(x_0)}$, x_0', and $\sqrt{R(x_0')}$. Thus the meaning of $\sqrt{R(x')}$ is uniquely determined as soon as the values of $\sqrt{R(x)}$, $\sqrt{R(x_0)}$, and $\sqrt{R(x_0')}$ have been fixed.

This is Euler's addition theorem.

Thus we have succeeded in reducing the general differential equation (1) to a simpler form, as well as in proving the fundamental property which is known after its discoverer under the name of Euler's *addition theorem*.

We must bear in mind that all the results which we have obtained so far are of a purely *formal* nature. What we have actually found can be summed up as follows. We have found one algebraic function of x and s which satisfies equation (39) *identically*, and another algebraic function of x and x' which satisfies equation (41) *identically*. But this does not, however, show us anything about the nature of the functional relation that connects the variables x and u that satisfy the original differential equation (1); we do not even know whether any such relationship exists at all. Neither do we know whether the solutions contained in the formulas (39) and (41) are the most general ones possible. It is true that we have designated them by the name of *complete* integrals but our conclusions are valid whether there exist other functions satisfying these equations or not.

The problem of finding the functional relationship between the variables x and u which satisfy the differential equation (1) and of representing this relationship analytically, can be looked at from two essentially different points of view.

We can regard x as the independent variable and u as the dependent one, and study the relationship from this starting point. Then we immediately find

$$u = \int_{x_0}^{x} \frac{dx}{\sqrt{R(x)}},$$

or, the desired relationship is obtained by equating u to a certain *elliptic integral of x*. This was the sole point of view of mathematicians before the time of Abel and Jacobi. It was not possible to get very far along this line since this elliptic integral was not well defined. For Euler and Legendre, as well as for Abel and Jacobi, the integral had a meaning only when the coefficients of $R(x)$ were real, the limits of integration were real and the variable of integration ran through a range of real values, from x_0 to x, not containing a root of $R(x) = 0$. It is true that the elliptic integral has a much wider region of definition, and it is also true that the integral yields the complete solution of the problem if it is taken in its most general sense. This, however.

cannot be seen without leaving the domain of elementary mathematics, and employing several general theorems from function theory, the proofs of which rest on ideas unfamiliar to the founders of the theory with which we are concerned. Thus the adoption of the second point of view in regard to the problem of integrating equation (1) became a measure of sweeping importance. Instead of regarding x as the independent variable and u as the dependent one, one can turn the question around and consider u as the independent variable and x as dependent. The introduction of this idea was destined to lead to a complete solution of the problem, and by means much more simple than required by the other method; as a matter of fact, it appeared that x was a much simpler function of u than u of x.

Before we pass over to the development of these ideas, let us consider the importance of the addition theorem when we apply the former or the latter point of view to the integration of our differential equation.

Under the assumptions developed above for the former point of view and under the further assumption that there is no real root of $R(x) = 0$ between any two of the quantities x_0, x, x_0', x', we find that

$$\int_{x_0}^{x} \frac{dx}{\sqrt{R(x)}} = \int_{x_0'}^{x'} \frac{dx}{\sqrt{R(x)}},$$

or

$$\int_{x_0}^{x} \frac{dx}{\sqrt{R(x)}} + \int_{x_0}^{x_0'} \frac{dx}{\sqrt{R(x)}} = \int_{x_0}^{x'} \frac{dx}{\sqrt{R(x)}},$$

satisfies equation (41). If, furthermore, we can show that there is only one relationship between the variables x' and x which satisfies such an equation and which permits the assignment of arbitrary initial values for the variables, then we obtain the theorem of Euler:

If the sum or the difference of two elliptic integrals of the form

$$\int \frac{dx}{\sqrt{R(x)}}$$

is equated to a third integral of the same form, and if further the lower limit of integration and the coefficients of the function under the radical are the same for all three integrals, then the upper limit of integration of the third integral

is an algebraic function of the two other upper limits, the common lower limit, and of the corresponding values of $\sqrt{R(x)}$.

On the other hand, let us start out from the latter point of view, and assume that, at least on certain continua of values of u and x in the neighborhood of $u = 0$ and $x = 0$, we can find a uniquely determined function

$$x = \varphi(u),$$

satisfying the differential equation

$$\frac{dx}{\sqrt{R(x)}} = du,$$

and such that

$$x_0 = \varphi(0).$$

Then the function

$$x' = \varphi(u + v)$$

satisfies the differential equation

$$\frac{dx'}{\sqrt{R(x')}} = du,$$

if v is determined in such a manner that

$$x_0' = \varphi(v)$$

and, moreover, v is so small that both u and $u + v$ belong to the values of u for which the function $\varphi(u)$ is defined. Then $\sqrt{R(x)} = \varphi'(u)$ and $\sqrt{R(x_0)} = \varphi(v)$, and the addition theorem reads:

If there exists one and only one function $x = \varphi(u)$ *which is uniquely defined on a certain continuous range of the variable* u *in the neighborhood of* $u = 0$, *satisfying the differential equation* (1), *and if* u, v *and* $u \pm v$ *are values belonging to the range mentioned, then* $\varphi(u \pm v)$ *is a rational function of* $\varphi(u)$, $\varphi(v)$ $\varphi'(u)$ *and* $\varphi'(v)$.

§ 2.

Abel's deduction of the general solution of the differential equation

$$\frac{dx}{\sqrt{R(x)}} = du. \tag{1}$$

$$R(x) = Ax^4 + 4Bx^3 + 6Cx^2 + 4B'x + A'. \tag{2}$$

We have seen that the efforts to deduce the general solution of the equation (1) from the *elliptic integral*

$$\int_a^x \frac{dx}{\sqrt{R(x)}} = u \tag{3}$$

must fail, since the integral has a definite meaning only in very special cases as long as we wish to avoid *the general theory of complex integration.* The new idea, namely, to consider x as a function of u instead of the older practice of regarding u as a function of x, must appear rather strange at first sight. The older way of looking at the problem has the appearance of being the direct way, and if that fails the converse method would be still less likely to lead anywhere.

Let us assume that the coefficients of $R(x)$ are real, that the lower limit of integration, the constant a, is real and that $R(a)$ is positive. Further let us determine which of the square roots of $R(a)$ is to be designated by $\sqrt{R(a)}$. Then the elliptic integral, u is a real, single-valued and continuous function of its upper limit, x, provided x is real and lies between a and either of the two real roots of $R(x) = 0$ which are nearest to a, and, further the variable of integration is real and runs over the same range. With the same assumptions, the upper limit x is a real, single-valued and continuous function of the integral u.

With Jacobi we call the function x of u which satisfies the differential equation (1) *an elliptic function of the argument u.* Thus we have obtained an element of the elliptic function from the elliptic integral corresponding to a certain special continuous range of real values of the argument.

Abel had the happy thought of trying to form the elliptic function in its totality out of this element and he succeeded in solving this problem, which looked very difficult on the surface, by developing *two* essentially new ideas which we shall expand further below.

First, however, we have to settle a preliminary question of great importance. Let us suppose that the meaning of the root $\sqrt{R(x)}$ has been fixed at the

point $x = a$. Suppose further that we have found an element of the elliptic function, $x = \varphi(u)$, satisfying the differential equation (1), which is defined in a certain neighborhood of $u = 0$ as a single-valued, analytic function of u and takes on the value $x = a$ for $u = 0$. Does there exist any other element of the elliptic function, $x = \psi(u)$, single-valued and analytic in u in the same region such that $a = \psi(0)$? Weierstrass has shown that the answer to this important question can easily be obtained from the transformation of equation (1) carried through in the preceding paragraph.

From equations (38) and (39) of § 1, we conclude that the differential equation

$$(4)\qquad \frac{ds}{\sqrt{S}} = \frac{dx}{\sqrt{R(x)}}$$

is satisfied identically by a relation of the form

$$(5)\qquad s = f\left[x, x_0, \sqrt{R(x)}, \sqrt{R(x_0)}\right],$$

where f is a certain rational function of x, x_0, $\sqrt{R(x)}$ and $\sqrt{R(x_0)}$. The particular square root of S which is denoted by $\sqrt{S}$ must be chosen in such a manner that equation (40), § 1, is satisfied, namely

$$(6)\qquad \sqrt{S} = f_1\left[x, x_0, \sqrt{R(x)}, \sqrt{R(x_0)}\right],$$

where f_1 is also a rational function of x, x_0, $\sqrt{R(x)}$, and $\sqrt{R(x_0)}$. If the equation (5) is differentiated with respect to x, we obtain

$$(7)\qquad \frac{\partial s}{\partial x} = \frac{f_1\left[x, x_0, \sqrt{R(x)}, \sqrt{R(x_0)}\right]}{\sqrt{R(x)}}.$$

If we permute x and x_0 in f and then differentiate with respect to x_0, we obtain

$$(8)\qquad \frac{f_1\left[x_0, x, \sqrt{R(x_0)}, \sqrt{R(x)}\right]}{\sqrt{R(x_0)}}.$$

The function f remains unchanged when x and x_0 are interchanged; f_1 on the other hand changes its sign by such a permutation but is otherwise unaffected. Thus we obtain

$$\frac{\partial s}{\partial x_0} = \frac{f_1\left[x, x_0, \sqrt{R(x)}, \sqrt{R(x_0)}\right]}{\sqrt{R(x_0)}}.$$

If in equation (5) we put $x = \varphi(u)$ and $x_0 = \psi(u)$ with $a = \varphi(0) = \psi(0)$ and $\sqrt{\varphi(0)} = \sqrt{\psi(0)}$ and, finally, differentiate with respect to u, we obtain

$$\text{(9)} \qquad \frac{ds}{du} = \frac{\partial s}{\partial x} \cdot \frac{dx}{du} + \frac{\partial s}{\partial x_0} \cdot \frac{dx_0}{du},$$

or, in view of (7) and (8),

$$\text{(10)} \qquad \frac{ds}{du} = 0.$$

Thus s is independent of u. But if $\varphi(u)$ and $\psi(u)$ actually are *different* functions of u, there must be some value of u within the common region of definition of the two functions for which they possess different values neither of which is infinite. At such a point s has a finite, perfectly definite value. On the one hand, s is a continuous function, on the other hand, a constant within the common region of definition of the two functions; thus s will keep the same *finite* value throughout this region. This, however, it not possible since at $u = 0$, where $x = x_0 = a$ and $\sqrt{R(x)} = \sqrt{R(x_0)} = \sqrt{R(a)}$, s becomes *infinite*, at least if a is *finite*. This remains true even if a should be *infinite* or, in other words, x and x_0 increase indefinitely when u approaches zero. It is true that $\sqrt{R(x)}$ does not remain finite when x becomes infinite; this is, however, the case with

$$\frac{\sqrt{R(x)}}{x^2} \quad \text{if only } A \neq 0.$$

Instead of setting $\sqrt{R(x)} = \sqrt{R(x_0)} = \sqrt{R(a)}$ as before, we lay down the condition

$$\lim_{x \to \infty} \frac{\sqrt{R(x)}}{x^2} = \lim_{x_0 \to \infty} \frac{\sqrt{R(x_0)}}{x_0^2}.$$

Thus we see that s becomes *infinite* like

$$\frac{2Ax^2x_0^2}{(x-x_0)^2}$$

when x and x_0 increase indefinitely. If, on the other hand, $A = 0$ then

$$\frac{\sqrt{R(x)}}{x\sqrt{x}}$$

approaches a definite limit provided the meaning of the radical $\sqrt{x}$ has been fixed. If we choose $\sqrt{x}$ and $\sqrt{x_0}$ in such a way that

$$\lim_{x\to\infty}\frac{\sqrt{R(x)}}{x\sqrt{x}} = \lim_{x_0\to\infty}\frac{\sqrt{R(x_0)}}{x_0\sqrt{x_0}},$$

then s becomes *infinite* like

$$2Bxx_0\left(\frac{\sqrt{x}+\sqrt{x_0}}{\sqrt{x}-\sqrt{x_0}}\right)^2$$

when x and x_0 increase indefinitely. Some doubt may arise whether this expression actually becomes infinite in the case when $\lim\frac{x_0}{x} = 1$. If in this case $\lim\frac{\sqrt{x_0}}{\sqrt{x}} = -1$, the expression would not necessarily become infinite. But $\lim\frac{\sqrt{x_0}}{\sqrt{x}}$ cannot equal -1; it must equal $+1$ when $\lim\frac{x_0}{x} = +1$, since

$$\lim\frac{x_0}{x} = \lim\frac{\frac{dx_0}{du}}{\frac{dx}{du}} = \lim\frac{\sqrt{R(x_0)}}{\sqrt{R(x)}} = \lim\frac{x_0\sqrt{x_0}}{x\sqrt{x}} = \lim\frac{\sqrt{x_0}}{\sqrt{x}}.$$

Thus, s will always become infinite at $u = 0$, no matter whether x and x_0 approach a finite limit a or increase indefinitely when u approaches zero. Hence, it is impossible that $\varphi(u)$ and $\psi(u)$ are different functions.

From this we conclude: *If it is possible to form a single-valued and analytic function $x = \varphi(u)$, defined on a certain neighborhood of $u = 0$, which satisfies the differential equation* (1), *and which either takes on a finite definite value a at $u = 0$, or becomes infinite when u approaches zero, then this function is the only function of u which has these properties within the given region. The meaning of the radical in equation* (1) *is defined in the following manner. If a is a finite quantity, not a root of $R(x) = 0$, then the meaning of $\sqrt{R(a)}$ is defined. If x should increase indefinitely when u approaches zero, we determine what meaning to give to*

$$\sqrt{A} = \lim_{x \to \infty} \frac{\sqrt{R(x)}}{x^2}.$$

It is not necessary that the continuum of values of u on which the function $\varphi(u)$ is defined forms a *two-dimensional manifold*, a *one-dimensional manifold* would do just as well. The term *analytic* is then to be understood to imply that the function $\varphi(u)$ is *continuous* within the region in question and its first derivative or the limit of $\dfrac{\Delta\varphi(u)}{\Delta u}$ is independent of the manner in which Δu approaches zero, provided $u + \Delta u$ remains within the region.

We have found how to form an element of the *elliptic function* directly from the *elliptic integral*, certain conditions being fulfilled, which corresponds to a certain continuum of real values of the argument. Now we also know that this function is the only element which corresponds to these values. *Thus the elliptic function is a single-valued and analytic function of its argument within the region in question.*

We now proceed to find out whether there also exists an elliptic function for other values of the argument, and we shall try to form this function out of the element obtained from the elliptic integral.

Comparing the theorem just obtained with the result of the preceding paragraph, we find that the study of the general elliptic function which for $u = 0$ has an arbitrary finite or infinite value a and satisfies the general equation (1), can be reduced to the study of the special elliptic function with the initial value ∞ which satisfies the equation

$$(11) \qquad \frac{ds}{\sqrt{S}} = du, \qquad S = 4s^3 - g_2 s - g_3.$$

As a matter of fact, having found a function $x = \varphi(u)$ with $a = \varphi(0)$ which satisfies equation (1), then (11) will be satisfied by another function $s = p(u)$

with $p(0) = \infty$, which is a certain rational function of x, $\sqrt{R(x)}$, a and $\sqrt{R(a)}$. In the same manner, having found a function $s = p(u)$, which satisfies equation (11) and for which $p(0) = \infty$, we can always form a certain function, $x = \varphi(u)$, rational in s, $\sqrt{S}$, a and $\sqrt{R(a)}$ which satisfies equation (1) and $\varphi(0) = a$. With the former assumption, let $\varphi(u)$ be an analytic function of u in a certain neighborhood of $u = 0$; then $p(u)$ will also be a single-valued and analytic function of u within the same region and the equation (11) has no other solution which becomes infinite at $u = 0$. Likewise with the latter assumption, that $p(u)$ is analytic within a certain neighborhood of $u = 0$, $\varphi(u)$ will also be a single-valued and analytic function of u within the same region, and this is the *only* analytic function of u which satisfies equation (1) for these values of u and which takes on the value a at $u = 0$. Thus we can restrict ourselves to finding the elliptic function which satisfies equation (11). If we succeed in determining this function for all values of the argument then we have also complete information about the elliptic function which satisfies equation (1), and this function will also be defined for all finite values of the argument.

In the following we shall denote the elliptic function which satisfies (11) and becomes infinite when u approaches zero by the Weierstrassian notation

$$(12) \qquad s = \wp(u) = \wp(u \mid g_2, g_3).$$

In order that $\wp(u)$ shall not degenerate into an exponential or a trigonometric function it is necessary that

$$g_2^3 - 27 g_3^2 \neq 0,$$

or, in other words, the three roots of $S = 0$, which we shall denote by e_1, e_2, e_3, shall all be *unequal*. This is indeed no *essential* restriction of our problem, since the case in which two roots or all three are equal is contained as a limiting case in our investigation. *We introduce an essential restriction, however, by assuming from now on that all the roots of $S = 0$ are real, or, in other words, that the quantities g_2 and g_3 are real and, moreover, $g_2^3 - 27 g_3^2 > 0$.* We choose the subscripts in such a way that

$$(13) \qquad e_1 > e_2 > e_3.$$

We have

$$e_1 + e_2 + e_3 = 0,$$

(14) $$4(e_1 e_2 + e_1 e_3 + e_2 e_3) = -g_2,$$

$$4 e_1 e_2 e_3 = g_3.$$

We further agree that $\sqrt{S}$ shall be positive for $s = +\infty$.

Let us form the elliptic integral

(15) $$\int_{+\infty}^{s} -\frac{ds}{\sqrt{S}} = u,$$

assuming the variable of integration to run through *real* values from $+\infty$ to e_1. Then u is a real, single-valued and analytic function of s for all values of s in the indicated range, and u increases continuously from the value zero to the real and positive value

(16) $$\int_{+\infty}^{e_1} -\frac{ds}{\sqrt{S}} = \int_{e_1}^{+\infty} \frac{ds}{\sqrt{S}} = \omega$$

when s decreases from $+\infty$ to e_1. With the same assumptions, s is a real, positive, single-valued and analytic function of u, which decreases from $+\infty$ to e_1, when u increases from zero to ω. Thus s is an element of the desired solution of equation (11), and from the definite integral (15) we have obtained an element of the function $\wp(u)$ corresponding to real and positive values of u between zero and ω. At the end-points of the range we have

(17) $$\wp(0) = +\infty, \quad \wp(\omega) = e_1.$$

We must now answer the question as to whether the function $\wp(u)$ exists for every finite value of u and see how to form the function in its totality from the known element. First we shall show how to form another element which corresponds to *purely imaginary* values of the variable between zero and a certain *purely imaginary* number ω_1. From the two elements obtained in this manner, we can immediately form two new elements corresponding to *real* values of the argument between zero and $-\omega$ and to purely imaginary values between zero and $-\omega_1$. The definition of the function $\wp(u)$ being obtained

for all *real* values of u between $-\omega$ and $+\omega$ and for all *purely imaginary* values between $-\omega_1$ and $+\omega_1$, we proceed to show how the function can be defined for all values of u by means of the addition theorem.

The first of these problems we solve by introducing one of the two fundamental ideas which underlie the theory of Abel. This idea is the simple observation that equation (11) can be written

$$(18) \qquad -\frac{d(c^3 s)}{\sqrt{c^6}\sqrt{4s^3 - g_2 s - g_3}} = du$$

where c denotes an arbitrary constant and $\sqrt{c^6} = c^3$. Consequently

$$(19) \qquad -\frac{ds_1}{\sqrt{4s_1^3 - c^4 g_2 s_1 - c^6 g_3}} = -\frac{ds_1}{\sqrt{S_1}} = du_1,$$

if we put

$$(20) \qquad s_1 = c^2 s, \qquad u_1 = \frac{u}{c}.$$

It is to be observed that the meaning of the radical $\sqrt{S_1}$ is uniquely defined only for real values of s between $+\infty$ and e_1.

Let us assume that the three roots, $\bar{e}_1$, $\bar{e}_2$, $\bar{e}_3$, of the equation

$$(21) \qquad 4s_1^3 - c^4 g_2 s_1 - c^6 g_3 = 0$$

are real, or, in other words, that the quantities $c^4 g_2$ and $c^6 g_3$ are real and, moreover,

$$(22) \qquad (c^4 g_2)^3 - 27(c^6 g_3)^2 = c^{12}(g_2^3 - 27 g_3^2) > 0.$$

Then, for all values of u_1 between zero and the quantity

$$(23) \qquad \bar{\omega} = \int_{\bar{e}_1}^{+\infty} \frac{ds}{\sqrt{4s^3 - c^4 g_2 s - c^6 g_3}}.$$

there exists a uniquely determined function of u_1, namely

$$(24) \qquad s_1 = \wp(u_1 \mid c^4 g_2, c^6 g_3)$$

which satisfies equation (19). The meaning of $\sqrt{S_1}$ has to be chosen in such a way that it will be positive when $s_1 = +\infty$. Since the expression (24) satisfies equation (19), equation (11) will be satisfied by

$$(25) \qquad s = \frac{1}{c^2}\wp\left(\frac{u}{c} \mid c^4 g_2,\, c^6 g_3\right)$$

for all real values of u/c between zero and $\bar{\omega}$. In order that the constants $c^4 g_2$ and $c^6 g_3$ shall satisfy the prescribed conditions it is necessary and sufficient that c^2 be a real quantity.

If we choose

$$(26) \qquad c^2 = -1, \quad c = i,$$

we find that equation (11) is satisfied by

$$(27) \qquad s = -\wp\left(\frac{u}{i} \mid g_2,\, -g_3\right)$$

for all *purely imaginary* values of u between zero and $\bar{\omega} i$. In order to determine $\bar{\omega}$ we have only to find the root $\bar{e}_1$. The three roots of equation (21) are $c^2 e_1$, $c^2 e_2$, and $c^2 e_3$. In view of (26), the greatest of these roots $\bar{e}_1$ equals $-e_3$. Thus we find

$$(28) \qquad \frac{\omega_1}{i} = \bar{\omega} = \int_{-e_3}^{+\infty} \frac{ds}{\sqrt{4s^3 - g_2 s + g_3}}$$

where ω_1 is *a purely imaginary quantity*. We have agreed upon denoting by $\wp(u \mid g_2, g_3)$ every function s of u which satisfies equation (11) and becomes infinite when u approaches zero. Hence, for purely imaginary values of u between zero and ω_1, we have

$$(29) \qquad \wp(u \mid g_2, g_3) = -\wp\left(\frac{u}{i} \mid g_2,\, -g_3\right)$$

with

$$(30) \qquad \wp(0 \mid g_2, g_3) = -\infty, \qquad \wp(\omega_1 \mid g_2, g_3) = e_3.$$

Thus we have formed a new element, corresponding to *purely imaginary* values of the argument between zero and ω_1, out of the previously known element of $\wp(u)$ which corresponds to *real* and *positive* values of u from zero to ω.

A repeated application of Abel's idea as contained in equations (18) and (19) will produce two new elements of the function $\wp(u)$ corresponding, one to real and negative values of u between $-\omega$ and zero, and the other to purely imaginary values between zero and $-\omega_1$. Instead of choosing ϵ^2 as in (26), set

$$\epsilon^2 = +1, \quad \epsilon = -1$$

whence

$$\wp(u \mid g_2, g_3) = \wp(-u \mid g_2, g_3) \tag{31}$$

which holds for all real and negative values of u between zero and $-\omega$ and for all purely imaginary values between zero and $-\omega_1$. Thus:

The equations (15), (27) and (31) determine the function $\wp(u)$ for all real values of u between $-\omega$ and $+\omega$ and also for all purely imaginary values between $-\omega_1$ and $+\omega_1$. For these values of the argument $\wp(u)$ is a real, single-valued and analytic function, finite everywhere excepting the point zero where

$$\wp(0) = \infty. \tag{32}$$

Further, $\wp(u)$ is an even function of u, that is to say, the relation

$$\wp(-u) = \wp(u) \tag{33}$$

holds throughout the region of definition. For the same values of u, the function $\wp'(u)$ is an odd function of u, thus

$$\wp'(-u) = -\wp'(u). \tag{34}$$

Finally

$$[\wp'(u)]^2 = 4[\wp(u)]^3 - g_2\wp(u) - g_3. \tag{35}$$

We have shown above that, once an element of $\wp(u)$ has been formed, it is possible to form a corresponding element of the general elliptic function, $\varphi(u)$, which satisfies equation (1) and takes on the prescribed value a at $u = 0$. Thus it is possible to define this function for all *real* values between $-\omega$ and $+\omega$ and for all *purely imaginary* values between $-\omega_1$ and $+\omega_1$.

Having developed the first of Abel's fundamental ideas on which his theory of elliptic functions is built, we proceed to the second leading idea. This will

enable us to express the function $\wp(u)$ and consequently also $\varphi(u)$ for all values of u in terms of the known elements. This requires a renewed study of the addition theorem.

Regarding s, x and x_0 as variables in equation (5) and differentiating with respect to the latter two, we find

$$ds = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial x_0} dx_0.$$

Thus the differential equation

$$-\frac{ds}{\sqrt{S}} = \frac{dx}{\sqrt{R(x)}} - \frac{dx_0}{\sqrt{R(x_0)}} \tag{36}$$

is identically satisfied by equation (5) in view of (6), (7) and (8). We also know how to form a rational function, x', of s and $\sqrt{S}$ which takes on the value a when s becomes infinite and which satisfies the equation

$$\frac{dx'}{\sqrt{R(x')}} = -\frac{ds}{\sqrt{S}}. \tag{37}$$

Introducing in this rational function the expressions for s and $\sqrt{S}$ in terms of x and x_0 as given by formulas (5) and (6), we find that x' is a certain rational function of x, $\sqrt{R(x)}$, x_0 and $\sqrt{R(x_0)}$, that is

$$x' = f_2\left[x, \sqrt{R(x)}, x_0, \sqrt{R(x_0)}\right]. \tag{38}$$

This relation (38) satisfies the differential equation

$$\frac{dx'}{\sqrt{R(x')}} = \frac{dx}{\sqrt{R(x)}} - \frac{dx_0}{\sqrt{R(x_0)}} \tag{39}$$

dentically. If in relation (38) we equate x to $\varphi(u)$ and x_0 to $\varphi(v)$, where u and v belong to the values of the argument for which the function φ is defined, we obtain

$$\frac{dx'}{\sqrt{R(x')}} = d(u-v). \tag{40}$$

The function which satisfies the differential equation (40) and takes on the value of a for $u-v=0$, we have denoted by $\varphi(u-v)$. Hence

$$x' = \varphi(u-v)$$

and, consequently, $\varphi(u-v)$ *is a certain rational function of* $\varphi(u)$, $\varphi'(u)$, $\varphi(v)$, $\varphi'(v)$. On the other hand, just as x' is a rational function of x $\sqrt{R(x)}$, x_0 and $\sqrt{R(x_0)}$, we know that x is a rational function of x', $\sqrt{R(x')}$, x_0 and $\sqrt{R(x_0)}$. Thus, replacing $u-v$ by u above we find that $\varphi(u+v)$ *is also a rational function of* $\varphi(u)$, $\varphi'(u)$, $\varphi(v)$ *and* $\varphi'(v)$. These two theorems together form the *addition theorem* for the function $\varphi(u)$. This proof is somewhat different from that given towards the end of the preceding paragraph. The reason for giving this different proof is essentially that it connects closely with Weierstrass' proof for the uniqueness of the solution of the differential equation (1); thus we do not need to step outside of the range of ideas within which we move in this parapraph. The given proof is as a matter of fact fundamentally the same as that given by Abel in his "*Recherches sur les fonctions elliptiques*".

The second fundamental idea of Abel's theory is the following. Using the first fundamental idea we can form $\varphi(u \pm v)$ as soon as $u \pm v$ belongs to the continuum of real numbers between $-\omega$ and $+\omega$ or purely imaginary values between $-\omega_1$ and $+\omega_1$. On the other hand, the addition theorem, or equation (38) and the analogous relation expressing x rationally in terms of x', $\sqrt{R(x')}$, x_0 and $\sqrt{R(x_0)}$, will define $\varphi(u \pm v)$ as soon as u and v themselves belong to these continua. Now it may very well happen that $u \pm v$ does not belong to these values though both u and v do so. Therefore the addition theorem gives us means to extend the region within which the elliptic function is defined beyond the limits reached so far. Repeating this process we can evidently carry this extension so far that the elliptic function becomes defined for all finite arguments.

We have already seen that it is enough to carry through this investigation for the function $\wp(u)$. To do this will be our next object. For $\wp(u)$ the equation (39) coincides with equation (36). Putting $x = s_1$ and $x_0 = s_0$ in this equation and in (5) we find that the equation

$$\text{(41)} \qquad -\frac{ds}{\sqrt{S}} = \frac{ds_1}{\sqrt{S_1}} - \frac{ds_0}{\sqrt{S_0}}$$

is satisfied identically by

$$\text{(42)} \qquad s = \frac{S(s_0) + \sqrt{S(s_1)}\,\sqrt{S(s_0)}}{2(s_1-s_0)^2} + \frac{S'(s_0)}{4(s_1-s_0)} + \tfrac{1}{24} S''(s_0)$$

or by the relation

$$(43)\qquad s = \frac{(s_1+s_0)(2s_1s_0-\tfrac{1}{2}g_2)-g_3+\sqrt{S(s_1)}\sqrt{S(s_0)}}{2(s_1-s_0)^2}$$

which is symmetric in s_1 and s_0. Substituting $s_1 = \wp(v)$ and $s_0 = \wp(u)$ in (43) and considering equations (33) and (34), we obtain the relation

$$(44)\quad \wp(u\pm v) = \frac{[\wp(u)+\wp(v)][2\wp(u)\wp(v)-\tfrac{1}{2}g_2]-g_3\mp\wp'(u)\wp'(v)}{2[\wp(u)-\wp(v)]^2}.$$

This relation holds whenever u and v are real numbers between $-\omega$ and $+\omega$ or purely imaginary numbers between $-\omega_1$ and $+\omega_1$. To begin with, we shall show how to define $\wp(u)$ for *all real* and *all purely imaginary* values of u, using the addition formula (44).

We observe that

$$(45)\qquad \wp(\omega) = e_1 \text{ and } \wp(\omega_1) = e_3$$

and

$$(46)\qquad [\wp'(u)]^2 = 4[\wp(u)-e_1][\wp(u)-e_2][\wp(u)-e_3]$$

and, consequently,

$$(47)\qquad \wp'(\omega) = 0 \text{ and } \wp'(\omega_1) = 0.$$

Thus if we equate v in (44) first to ω and then to ω_1, we obtain

$$(48)\qquad \wp(u\pm\omega) = \frac{[\wp(u)+\wp(\omega)][2\wp(u)\wp(\omega)-\tfrac{1}{2}g_2]-g_3}{2[\wp(u)-\wp(\omega)]^2},$$

$$(49)\qquad \wp(u\pm\omega_1) = \frac{[\wp(u)+\wp(\omega_1)][2\wp(u)\wp(\omega_1)-\tfrac{1}{2}g_2]-g_3}{2[\wp(u)-\wp(\omega_1)]^2}.$$

The right hand side of equation (48) has a perfectly definite meaning for all real values of u between $-\omega$ and $+\omega$. Hence if we let u increase from zero to $+\omega$ and use the plus sign on the left hand side, the function $\wp$ will be uniquely determined by (48) for all values of the argument between $+\omega$ and $+2\omega$. If we use the minus sign instead and let u decrease from zero to $-\omega$,

we get the function uniquely defined for all values of the argument between $-\omega$ and -2ω. The equation

$$\wp(-u) = \wp(u) \tag{50}$$

still holds for the new values of u. In this manner the right hand side of (48) has a uniquely defined meaning for real values of u between $+\omega$ and $+2\omega$ and between $-\omega$ and -2ω. Then this equation determines $\wp(u)$ also for real values of u between $+2\omega$ and $+3\omega$ and between -2ω and -3ω. By repeating the same process the function $\wp(u)$ will be defined for all *real* values of u. The equation (50) always holds. Using equation (49) in the same manner, we obtain the function $\wp(u)$ defined for all *purely imaginary* values of u. The equation (50) still holds. Finally, if in equation (44) we let u run through all real values and v through all purely imaginary values, the function $\wp(u)$ will be defined for all finite complex values of u, and the relation (50) continues to hold for all values of u. Thus the repeated application of the addition theorem enables us to define the function $\wp(u)$ for all finite values of the argument u.

One point may cause doubt: is the function so defined an *analytic* function, or, in other words, is $\lim \dfrac{\Delta\wp(u)}{\Delta u}$ independent of the way in which Δu tends towards zero? This difficulty is settled, however, by formula (40). Putting $u = \xi$ and $v = -i\eta$ where ξ and η are real numbers, formula (40) gives

$$\sqrt{S[\wp(\xi + i\eta)]} = \frac{d\wp(\xi + i\eta)}{d(\xi + i\eta)} = \lim \frac{\Delta\wp(\xi + i\eta)}{\Delta(\xi + i\eta)}.$$

Thus the derivative of $\wp(\xi + i\eta)$ is independent of the way in which $\Delta(\xi + i\eta)$ approaches zero.

Still another question remains to be settled. The function $\wp(u)$ is by definition that particular solution of equation (11) which becomes infinite at $u = 0$, and is uniquely characterized by that property. In order that the function formed by the method described above shall actually be equal to $\wp(u)$ it must also have the same property. We know that this function becomes positively infinite if the argument approaches zero through *real* values. Further we know that the function becomes negatively infinite if the argument approaches zero through *purely imaginary* values. We must also require that the function becomes infinite no matter how we let the variable approach zero. This is actually the case, as follows from the result at the beginning of this paragraph. If we substitute $s_1 = \wp(\xi)$ and $s_0 = \wp(i\eta)$ in equation (43), ξ and η being real

quantities, we obtain the value of our function for the argument $\xi + i\eta$. If we let ξ and η both approach zero, s_1 and s_0 increase indefinitely. The equation (43) is only a special case of equation (5) and we have seen that s increases indefinitely in this equation if we let x and x_0 increase indefinitely at the same time. Consequently s in equation (43) increases indefinitely at the same time as s_1 and s_0 become infinite; thus our function has the required property of becoming infinite wheneve $\xi + i\eta$ approaches zero, no matter how ξ and η approach zero.

Thus we have found that *the equation* (11) *is satisfied by a single-valued analytic function of u, $\wp(u)$ which increases indefinitely when u approaches zero. The way to form this function is uniquely defined. This function is the only analytic function of u which satisfies* (11) *and becomes infinite when u approaches zero. It is an even function of u and has, moreover, an addition theorem expressed by formula* (44).

The addition theorem gives some other properties of $\wp(u)$. From equations (48) and (49) it follows that

$$(51)\qquad \wp(u+\omega) = \wp(u-\omega), \qquad \wp(u+\omega_1) = \wp(u-\omega_1),$$

and consequently

$$(52)\qquad \wp(u+2\omega) = \wp(u), \qquad \wp(u+2\omega_1) = \wp(u).$$

By iteration of these formulae we find

$$(53)\qquad \wp(u+2m\omega+2m_1\omega_1) = \wp(u)$$

where m and m_1 denote positive or negative whole numbers.

Hence, $\wp(u)$ takes on the same value when u is increased by the quantity $2m\omega + 2m_1\omega_1$ where ω and ω_1 are determined by the formulae (16) *and* (28) *and m and m_1 denote positive or negative integers. $\wp(u)$ is a periodic function of u; since all the periods are formed from the periods 2ω and $2\omega_1$ by addition and substraction, we call $\wp(u)$ a doubly periodic function of u.*

We are now able to determine the points where our function becomes *infinite*. We have shown that

$$\wp(0) = \infty.$$

In view of the periodicity of the $\wp$-function it follows that

$$\wp(2m\omega + 2m_1\omega_1) = \infty. \tag{54}$$

It is not self-evident, however, that formula (54) gives *all* the points where $\wp(u)$ becomes infinite; therefore it will be necessary to investigate whether or not other infinities are possible. That *all* infinities are actually given by formula (54) can be proved as follows. It follows from formula (15) that $\wp(u)$ can not become infinite for any *real* value of u in the interval from zero to ω, the upper limit included. In view of the first of formulae (51) and the equation $\wp(-u) = \wp(u)$, we conclude that $\wp(u)$ does not become infinite for any real value u between ω and 2ω. Then by the first of formulae (52) and the equation $\wp(-u) = \wp(u)$ it follows that the only *real* values of u for which $\wp(u)$ becomes infinite are of the form $2m\omega$. By similar reasons it follows that the only *purely imaginary* values for which $\wp(u)$ becomes infinite are of the form $2m_1\omega_1$. It remains to settle the question whether or not $\wp(u)$ may become infinite for a value $\xi + i\eta$ where the real quantities ξ and η are different from zero, and, furthermore, the equalities $\xi = 2m\omega$ and $i\eta = 2m_1\omega_1$ do not hold simultaneously. However, we find that $\wp(u)$ can not become infinite for such a value in view of the way in which it has been formed by means of formula (44). Thus:

The function $\wp(u)$ becomes infinite when the argument u approaches one of the infinitely many values given by the formula

$$u = 2m\omega + 2m_1\omega_1.$$

The function can not become infinite for any other value of u.

The next question will be to determine *the mode* in which $\wp(u)$ becomes infinite at these points. Since $\wp(u)$ is a holomorphic function of u in any neighbourhood of one of the infinities, $u = 0$ say, $\wp(u)$ will presumably become infinite in such a manner that the quotient of the function and a certain power of $1/u$ approaches a definite constant value when u approaches zero. We can determine the limiting value of

$$u\,\wp(u) \tag{55}$$

since $\wp'(u)$ and all the following derivatives of $\wp(u)$ can be expressed in terms of $\wp(u)$ by means of formula (46). It is immediately shown that the limiting

value of (55) is ∞, which might be expected since $\wp(u)$ is an *even* function of u. Then we consider the limit of

$$(56) \qquad u^2 \wp(u)$$

which is easily shown to be $+1$ when u approaches zero. Consequently, in the neighbourhood of $u = 0$ we have

$$(57) \qquad \wp(u) = \frac{1}{u^2} + \frac{\varepsilon(u)}{u^2}$$

where $\varepsilon(u)$ vanishes with u. In order to determine $\varepsilon(u)$ a little closer we look for the limit of

$$\frac{u^2 \wp(u) - 1}{u^2} = \frac{\varepsilon(u)}{u^2}.$$

It can be shown without any difficulty that this limit is zero as well as the limit of

$$\frac{u^2 \wp(u) - 1}{u^3},$$

whereas the function

$$\frac{u^2 \wp(u) - 1}{u^4}$$

approaches the limit

$$\frac{g_2}{20}$$

when u approaches zero. Thus in the neighbourhood of $u = 0$:

$$(58) \qquad \wp(u) = \frac{1}{u^2} + \left[\frac{g_2}{20} + \delta(u)\right] u^2$$

where $\delta(u)$ vanishes with u. This function $\delta(u)$ moreover, is *finite* for all values of u the *real* part of which lies between -2ω and $+2\omega$ and the *purely imaginary* part of which lies between $-2\omega_1$ and $+2\omega_1$. From the way in which formula (58) has been obtained and in view of the fact that

$\wp(u)$ as well as all its derivatives are doubly periodic functions, it follows that $\wp(u)$ can be expressed by the formula

$$\wp(u + 2m\omega + 2m_1\omega_1) = \frac{1}{u^2} + \left[\frac{g_2}{20} + \delta(u)\right]u^2 \tag{59}$$

in the neighbourhood of some other infinity $2m\omega + 2m_1\omega_1$. Having settled *the question as to the points in which $\wp(u)$ becomes infinite and the way in which it becomes infinite*, we shall now derive certain formulae from the addition theorem which will help us to answer the question whether the function $\wp(u)$ will take on a given value at other points than those given by formula (53).

Substituting $s_1 = \wp(u)$ and successively $s_0 = \wp(\omega)$ and $s_0 = \wp(\omega_1)$ in the formula (42) we obtain, in view of formulae (45), (46) and (47),

$$\wp(u \pm \omega) - e_1 = \frac{(e_1 - e_2)(e_1 - e_3)}{\wp(u) - e_1}, \tag{60}$$

$$\wp(u \pm \omega_1) - e_3 = \frac{(e_3 - e_1)(e_3 - e_2)}{\wp(u) - e_2}. \tag{61}$$

From these formulae it follows immediately that

$$\wp(\omega \pm \omega_1) = e_2. \tag{62}$$

Consequently, putting $s_1 = \wp(u)$ and $s_0 = \wp(\omega \pm \omega_1)$ in formula (42) we find

$$\wp(u \pm \omega \pm \omega_1) = \frac{(e_2 - e_1)(e_2 - e_3)}{\wp(u) - e_2}. \tag{63}$$

In view of (46) and (62)

$$\wp'(\omega \pm \omega_1) = 0. \tag{64}$$

Using the theorem of addition, formula (44), we obtain by adding and subtracting $\wp(u + v)$ and $\wp(u - v)$

$$(65)\qquad \wp(u+v)+\wp(u-v)=\frac{[\wp(u)+\wp(v)]\,[2\wp(u)\wp(v)-\tfrac{1}{2}g_2]-g_3}{2\,[\wp(u)-\wp(v)]^2},$$

$$(66)\qquad \wp(u+v)-\wp(u-v)=-\frac{\wp'(u)\,\wp'(v)}{[\wp(u)-\wp(v)]^2}.$$

For what values of u'' will

$$(67)\qquad \wp(u'')=\wp(u')?$$

This question is easily answered with the help of the preceding system of formulae. Putting

$$(68)\qquad u''=u+v,\qquad u'=u-v$$

in (66) we find

$$(69)\qquad \wp(u'')-\wp(u')=-\frac{\wp'\left(\frac{u''+u'}{2}\right)\wp'\left(\frac{u''-u'}{2}\right)}{\left[\wp\left(\frac{u''+u'}{2}\right)-\wp\left(\frac{u''-u'}{2}\right)\right]^2}.$$

Thus the equation (67) can hold if

$$(70)\qquad \wp\left(\frac{u''+u'}{2}\right)=\infty,$$

or if

$$(71)\qquad \wp\left(\frac{u''-u'}{2}\right)=\infty,$$

or if

$$(72)\qquad \wp'\left(\frac{u''+u'}{2}\right)=0,$$

or, finally, if

$$(73)\qquad \wp'\left(\frac{u''-u'}{2}\right)=0,$$

and only if one of these equations is fulfilled. Equation (70) will hold if and only if

$$u''=-u'+2(2m\,\omega+2m_1\,\omega_1).$$

Likewise, equation (71) holds if and only if

$$u'' = u' + 2(2m\omega + 2m_1\omega_1).$$

In order to investigate the meaning of equations (72) and (73) it is necessary to know at what points $\wp'(u)$ will vanish. The equation (46) tells us that $\wp'(u)$ will be zero if and only if $\wp(u)$ takes on one of the values e_1, e_2 or e_3. Since we know all the points where $\wp(u)$ becomes infinite, the formulae (60), (61) and (63) give us means of ascertaining all the points where $\wp(u)$ takes on one of these values. *We find*

$$\wp[(2m+1)\omega + 2m_1\omega_1] = e_1,$$

$$\wp[(2m+1)\omega + (2m_1+1)\omega_1] = e_2, \tag{74}$$

$$\wp[2m\omega + (2m_1+1)\omega_1] = e_3.$$

and consequently

$$\wp'(m\omega + m_1\omega_1) = 0, \tag{75}$$

where formulae (74) *give us all the values at which* $\wp(u)$ *takes on one of the values* e_1, e_2 *or* e_3, *and formula* (75) *gives all the zeros of* $\wp'(u)$. This implies that equation (72) will hold if and only if

$$u'' = -u' + 2m\omega + 2m_1\omega_1,$$

and equation (73) if and only if

$$u'' = u' + 2m\omega + 2m_1\omega_1.$$

Thus:

The function $\wp(u)$ *is a doubly periodic even function of* u *which takes on the same value at all the points which are expressed by the formula*

$$\pm u + 2m\omega + 2m_1\omega_1$$

but has a value different from this at all other points.

Strictly speaking we have solved the problem which was proposed at the beginning of this paragraph. We have shown that it is always possible to find a function $\varphi(u)$ which satisfies differential equation (1) and which takes

on the arbitrary value a at the point $u = 0$. All we had to do was to form a function $\wp(u)$ which satisfies the equation (11) and becomes infinite when u tends towards zero. $\varphi(u)$ was a certain rational function of $\wp(u)$ and $\wp'(u)$. We have been able to express this function $\wp(u)$ for every finite value of u and we have also found some of the most characteristic properties of this function. Further, we have seen that $\varphi(u)$ shares several of these properties, that $\varphi(u)$ as well as $\wp(u)$ is a single-valued and analytic function of u, and that $\varphi(u)$ also has an addition theorem. Since $\wp(u)$ and consequently also $\wp'(u)$ are doubly periodic functions of u, it follows that $\varphi(u)$ is a doubly periodic function with the same periods $2\,\omega$ and $2\,\omega_1$.

However, the elliptic functions would not be of such an immense importance to mathematics as they really are, if they did not possess still further another fundamental property, the discovery of which is also due to the genius of Abel. It is true that we know how to find the value of $\wp(u)$ and thence the value of $\varphi(u)$ for every finite value of u, but the method was different for different forms of the argument u and thus lacked the real mathematical character of simplicity and lucidity. This deficiency would be remedied if we could find a way of forming the function $\wp(u)$ which involved applying the same analytical operation for all values of the argument u. Owing to the analogy with the exponential and the trigonometric functions where uniform laws of formation exist, one feels justified in presuming the existence of such a law also in the present case. The exponential and trigonometric functions are expressible in terms of power series convergent everywhere. Such an expansion is not possible for the function $\wp(u)$ with its infinitely many poles, but there is perhaps some other kind of a series, the different terms of which are not simple powers of u, which converges everywhere towards the function except at the given poles. This is actually the case; $\wp(u)$ is expressible by such a general analytic expression which can be derived by a systematic continuation of the study of the addition theorem. This is the problem to be studied in the following paragraph.

§ 3.

Deduction, according to Abel, of the general analytic expression for the function $\wp(u)$.

The addition theorem, formula (44), tells us, that $\wp(u+v)$ is a rational function of $\wp(u)$, $\wp(v)$, $\wp'(u)$ and $\wp'(v)$. By iterated use of this formula we find that $\wp(u_1 + u_2 + \cdots + u_n)$ is a rational function of $\wp(u_1)$, $\wp(u_2), \ldots, \wp(u_n)$ and of $\wp'(u_1)$, $\wp'(u_2), \ldots, \wp'(u_n)$. Thus if we put

$$u_1 = u_2 \cdots = u_n = u$$

we find that $\wp(nu)$ is a rational function of $\wp(u)$ and $\wp'(u)$. Further, since $[\wp'(u)]^2$ is an integral rational function of $\wp(u)$, we can express $\wp(nu)$ in terms of $\wp(u)$ and $\wp'(u)$ by means of a formula of the type

$$\wp(nu) = R[\wp(u)] + \wp'(u)\, R_1[\wp(u)], \tag{1}$$

where R and R_1 are rational function of $\wp(u)$. In this formula R_1 must be identically zero, since $\wp(u)$ is an *even* function of u and consequently $\wp'(u)$ an *odd* function. Hence we have

$$\wp(nu) = R[\wp(u)] = \frac{G[\wp(u)]}{G_1[\wp(u)]}, \tag{2}$$

where G and G_1 denote polynomials in $\wp(u)$ which are relatively prime. If in (2) we replace nu by u we obtain

$$G_1\left[\wp\left(\frac{u}{n}\right)\right]\wp(u) = G\left[\wp\left(\frac{u}{n}\right)\right]. \tag{3}$$

To actually obtain and discuss the equation (3) by algebraic methods would offer considerable difficulties. However, if we call to our aid the theorems concerning the function $\wp(u)$ obtained in the preceding paragraph, it becomes an easy matter to carry through such an investigation.

Let us put

$$y = \wp(u), \qquad x = \wp\left(\frac{u}{n}\right); \tag{4}$$

then (3) becomes

$$G_1(x)\, y - G(x) = 0. \tag{5}$$

In view of the way in which equations (3) and (5) were obtained we have

$$-\frac{dy}{\sqrt{S(y)}} = -n\frac{dx}{\sqrt{S(x)}}$$

identically, and, moreover, we know that x and y become infinite simultaneously. If we consider a value of y obtained from the first of equations (4) and put

$$dv = -\frac{dy}{\sqrt{S(y)}} \tag{6}$$

then all the values of v which satisfy equation (6) for this particular value of y are contained in the formula

$$v = \pm u + 2m\omega + 2m_1\omega_1. \tag{7}$$

Assuming $y = \wp(u)$, all the roots of equation (5) are contained in the expression

$$x = \wp\left(\frac{\pm u + 2m\omega + 2m_1\omega_1}{n}\right), \tag{8}$$

in which m and m_1 are arbitrary positive or negative integers. If we assume n to be an *odd* number we notice immediately that all the values which the expression (8) is capable of taking on are obtained by letting m and m_1 separately take on all integral values from $-\frac{n-1}{2}$ to $+\frac{n-1}{2}$, including the limits, in the expression

$$\wp\left(\frac{u + 2m\omega + 2m_1\omega_1}{n}\right). \tag{9}$$

We also notice that all these values are actually different. Their number is n^2. In the same manner, if n is an *even* number we obtain all the different values of expression (8) by letting m and m_1 in (9) separately run through all the integers from $-\frac{n-2}{2}$ to $\frac{n}{2}$ including the limits. All these values are mutually different and are n^2 in number. In each case the equation (5) has n^2 different roots. All these roots are necessarily simple roots, otherwise

$$G_1(x)y - G(x) = 0, \qquad G_1'(x)y - G'(x) = 0,$$

simultaneously and independently of u, whence it would follow that

$$G_1G' - GG_1' = 0$$

or

$$\frac{d}{dx}\left(\frac{G}{G_1}\right) = 0$$

independently of u. Hence G/G_1 would have to be independent of x, and further y would have a constant value independent of u which, as we know, is not the case.

The equation

$$G_1(x)y - G(x) = 0$$

*is of degree n^2 with respect to x. Its n^2 roots, which can be obtained from formula (5) in the manner indicated, are all simple roots, at any rate if we restrict ourselves to such values of y which are obtained from the equation $y = \wp(u)$.**

It is well known that the symmetric functions of the roots of equation (5) are rational functions of the coefficients of the equation, and hence rational functions of $\wp(u)$. The sum of the roots is the simplest of these symmetric functions. This function is equal to the coefficient of the (n^2-1)st power of x if the coefficient of the n^2-th power is taken to be $+1$. Hence we have

$$(10) \qquad A + B\wp(u) = \sum_m \sum_{m_1} \wp\left(\frac{u + 2m\omega + 2m_1\omega_1}{n}\right),$$

where A and B do not depend upon u. In order to determine these constants we divide both sides of the equality by $\wp\left(\frac{u}{n}\right)$ and let u tend towards zero. The limit of the right hand side is $+1$, whence we obtain

$$\lim_{u \to 0} \frac{B\wp(u)}{\wp\left(\frac{u}{n}\right)} = +1,$$

and in view of equation (58) in § 2

$$(11) \qquad B = n^2.$$

Then we subtract $\wp\left(\frac{u}{n}\right)$ from both sides of equation (10) and let u tend towards zero. On account of equation (58) in § 2 we have

$$(12) \qquad \lim\left\{n^2\wp(u) - \wp\left(\frac{u}{n}\right)\right\} = 0.$$

* This restriction is necessary since it is by no means self-evident that y takes on *all* values when u runs through *all* values. It is true that this question can be settled by using the principles on which our theory is founded, but to carry out this investigation would take us too far from our main subject.

whence we obtain the following value of A, namely

$$(13) \qquad A = \sum_{m, m_1}{}' \wp\left(\frac{2m\omega + 2m_1\omega_1}{n}\right).$$

The primed summation sign denotes that the combination $m = 0$, $m_1 = 0$ has to be left out when we sum for m and m_1. If n is an *odd* number, m and m_1 take on all integral values between $-\frac{n-1}{2}$ and $+\frac{n-1}{2}$. If n is an *even* number all the integral values between $-\frac{n-2}{2}$ and $+\frac{n}{2}$ are taken on instead. Thus the equation (10) is carried over into the remarkable equality

$$(14) \quad n^2 \wp(u) = \wp\left(\frac{u}{n}\right) + \sum_{m, m_1}{}' \left\{\wp\left(\frac{u + 2m\omega + 2m_1\omega_1}{n}\right) - \wp\left(\frac{2m\omega + 2m_1\omega_1}{n}\right)\right\}.$$

This equality enables us to express $\wp(nu)$ as a linear function of n^2 functions of the form

$$\wp\left(u + \frac{2m\omega + 2m_1\omega_1}{n}\right).$$

This equality becomes still more remarkable if we let the arbitrarily chosen positive integer n increase indefinitely. The limiting value which we obtain in this manner is exactly the desired general analytic expression for the function $\wp(u)$.

If u and v are two different quantities, such that $u^2 \neq v^2$, formula (58) in § 2 yields

$$(15) \qquad \wp\left(\frac{u}{n}\right) - \wp\left(\frac{v}{n}\right) = \frac{n^2}{u^2} - \frac{n^2}{v^2} + g\,\frac{u^2 - v^2}{n^2},$$

provided n is sufficiently large. We have to choose n so large that the real parts of $\frac{u}{n}$ and $\frac{v}{n}$ belong to the interior of the interval $(-2\omega, +2\omega)$ and the imaginary parts belong to the interior of the interval $\left(-2\frac{\omega_1}{i}, +2\frac{\omega_1}{i}\right)$. Here g stands for a *finite* quantity which tends towards the value $\frac{g_2}{20}$ when n increases indefinitely and the absolute value of which is bounded.

If in formula (15) we substitute $u + 2m\omega + 2m_1\omega_1$ for u and $2m\omega + 2m_1\omega_1$ for v and use formula (15) on the equality (14), the latter becomes

$$(16)\quad \begin{aligned} \wp(u) = \frac{1}{u^2} + \sum_{m,m_1}{}' \left\{ \frac{1}{(u + 2m\omega + 2m_1\omega_1)^2} - \frac{1}{(2m\omega + 2m_1\omega_1)^2} \right\} \\ + \frac{1}{n^4} \sum_{m,m_1} g_{m,m_1}[u^2 + 2(2m\omega + 2m_1\omega_1)u]. \end{aligned}$$

The meaning of g_{m,m_1} is self-evident. Let us put

$$(17)\quad S_{0,n} = \frac{1}{u^2} + \sum_{m,m_1}{}' \left\{ \frac{1}{(u + 2m\omega + 2m_1\omega_1)^2} - \frac{1}{(2m\omega + 2m_1\omega_1)^2} \right\},$$

$$(18)\quad S_{1,n} = \frac{u^2}{n^4} \sum_{m,m_1}{}' g_{m,m_1},$$

$$(19)\quad S_{2,n} = \frac{2u}{n^4} \sum_{m,m_1}{}' g_{m,m_1}(2m\omega + 2m_1\omega_1),$$

and consequently

$$\wp(u) = S_{0,n} + S_{1,n} + S_{2,n}.$$

Let us further put

$$(20)\quad S = \frac{1}{u^2} + \sum_{m=-\infty}^{+\infty} \sum_{m_1=-\infty}^{+\infty}{}' \left\{ \frac{1}{(u + 2m\omega + 2m_1\omega_1)^2} - \frac{1}{(2m\omega + 2m_1\omega_1)^2} \right\}.$$

Now we proceed to show that the sum $S_{1,n} + S_{2,n}$ can be made less in absolute value than any preassigned positive number by increasing the value of n and consequently that the difference

$$\wp(u) - S_{0,n}$$

can be made to approach zero simply by increasing the number of terms in the sum $S_{0,n}$. From this it follows that $\wp(u)$ can be expressed by the series S for every value of u, and the equation

$$(21)\quad \wp(u) = S$$

will always hold. We shall subject the series (20) to a somewhat closer inspection.

With Weierstrass we denote the absolute value of x by $|x|$.

Denoting by G a finite positive quantity which is greater than all the quantities $|g_{m,m_1}|$ we have

$$|S_{1,n}| < G \cdot |u|^2 \cdot \frac{1}{n^2}, \tag{22}$$

and

$$|S_{2,n}| < 2 \cdot G \cdot |u| \cdot |2\omega + 2\omega_1| \cdot \frac{1}{n}. \tag{23}$$

From this it follows immediately that by increasing the value of n sufficiently, we can make $|S_{1,n} + S_{2,n}|$ less than any preassigned positive number.

Thus the equation (21) holds and the sum S gives a general analytical expression for the function $\wp(u)$ which holds for every finite value of u.

In order to investigate the sum S we assume

$$|u| \leq \varrho \tag{24}$$

where ϱ is an arbitrarily chosen positive number. Then we separate the sum S in two parts

$$S = S' + S''. \tag{25}$$

The letter S' denotes the sum of all the terms for which

$$|2m\omega + 2m_1\omega_1| < \varrho. \tag{26}$$

Thus this sum contains only a finite number of terms; this number changes with the size of ϱ but is otherwise always the same. The letter S'' denotes the sum of all the remaining terms for which

$$|2m\omega + 2m_1\omega_1| \geq \varrho. \tag{27}$$

The difference

$$\wp(u) - S'$$

is always *finite* when u lies in the region (24), since $\wp(u)$ becomes infinite only at the points $u = 2m\omega + 2m_1\omega_1$ and there is always a term in S' such that the difference of $\wp(u)$ and this term tends towards zero when u approaches the point in question. This is the only term in S' which becomes infinite at the point. On the other hand, S' does not become infinite for any other value of u than $u = 2m\omega + 2m_1\omega_1$. For such a value of u *one and only one* term becomes infinite, but the difference of $\wp(u)$ and the term in question approaches zero. Every term in S'' is *finite* for the values of u which belong to the region (24).

We shall show that the series S'' is *uniformly convergent* when u lies in the region determined by formula (24). This is obvious from the way in which the series S and S'' have been generated; it is also obvious that we can arrange the terms in S'' in an arbitrary manner. As a matter of fact, we can make the difference $\wp(u) - S_{0,n}$ as small as we please by increasing the number n. Let S_0^r stand for a sum with r terms and $S_0^{r+r_1}$ for one with $r+r_1$ terms. Then the difference

$$S_0^{r+r_1} - S_0^r$$

can be made less than an arbitrarily small positive number, δ, and we can always find an r such that

$$S_0^{r+r_1} - S_0^r < \delta$$

for every value of u in the region (24). But this is exactly what is needed in order to prove that *the series S'' is uniformly convergent.*

It is also an easy matter to show that *the series formed by the absolute values of the terms of S'' is uniformly convergent,* from which it follows that we have the right to rearrange the order of the terms in S''. Let us take the general term in the series S'', namely

$$(28)\quad \begin{aligned} &\frac{1}{(u+2m\omega+2m_1\omega_1)^2} - \frac{1}{(2m\omega+2m_1\omega_1)^2} \\ &= -\frac{1}{(2m\omega+2m_1\omega_1)^3}\cdot\frac{2u+\dfrac{u^2}{2m\omega+2m_1\omega_1}}{\left[1+\dfrac{u}{2m\omega+2m_1\omega_1}\right]^2}.\end{aligned}$$

In view of (24) and (27) the quantity

$$(29)\qquad \frac{2u + \dfrac{u^2}{2m\omega + 2m_1\omega_1}}{\left[1 + \dfrac{u}{2m\omega + 2m_1\omega_1}\right]^2}$$

is bounded when $|u| \leqq \varrho$. Thus *the series formed by the absolute values of the terms of the series S'' is uniformly convergent in* $|u| \leqq \varrho$, if we can show the *absolute convergence* of the series whose general term is

$$(30)\qquad \frac{1}{(2m\omega + 2m_1\omega_1)^3}.$$

In order to investigate* this latter series we order the terms in groups such that to the group of order μ belong all the terms for which

$$m = \pm\mu \text{ and } m_1 = 0, \pm 1, \ldots, \pm\mu,$$

and all the terms for which

$$m_1 = \pm\mu \text{ and } m = 0, \pm 1, \ldots, \pm(\mu - 1),$$

where μ denotes a positive integer. Let us denote by w the smaller of the two numbers $|2\omega|$ and $|2\omega_1|$, then we obtain

$$(31)\qquad \sum{}' \frac{1}{|2m\omega + 2m_1\omega_1|^3} < \sum 8\mu \frac{1}{(\mu w)^3} = \frac{8}{w^3}\sum\frac{1}{\mu^2}.$$

The series $\sum\frac{1}{\mu^2}$ being convergent we have shown that *the series formed by the absolute values of the terms of the series S'' is uniformly convergent* in the region given by formula (24).

* This method was used by Weierstrass in his lectures for the purpose of investigating the general series

$$\sum{}' \frac{1}{(2m\omega + 2m_1\omega_1)^\lambda}.$$

Finally we indicate how the passage from the Weierstrassian $\wp$-function to the Weierstrassian σ-function can be effected. We integrate the series (20) twice term-wise noticing that at the first integration the integral of the first term, $1/u^2$, is taken to be $-1/u$ and at the second integration $-\log u$, and in integrating the other terms care is taken to choose the constants of integration in such a way that each term vanishes with u. The series obtained after two integrations we take with opposite sign and use as an exponent for the number e, thus forming

$$(32)\quad \sigma(u) = u \prod_{m,m_1}{}' \left\{ \left(1 + \frac{u}{2m\omega + 2m_1\omega_1}\right) e^{-\frac{u}{2m\omega + 2m_1\omega_1} + \frac{1}{2}\left(\frac{u}{2m\omega + 2m_1\omega_1}\right)^2} \right\}.$$

$$(33)\quad \begin{aligned} \frac{\sigma'(u)}{\sigma(u)} = \frac{1}{u} + \sum_{m,m_1}{}' \Bigg\{ &\frac{1}{u + 2m\omega + 2m_1\omega_1} \\ &- \frac{1}{2m\omega + 2m_1\omega_1} + \frac{u}{(2m\omega + 2m_1\omega_1)^2} \Bigg\}. \end{aligned}$$

$$(34)\quad \begin{aligned} -\frac{d}{du}\left[\frac{\sigma'(u)}{\sigma(u)}\right] &= \wp(u) \\ &= \frac{1}{u^2} + \sum_{m,m_1}{}' \left\{ \frac{1}{(u + 2m\omega + 2m_1\omega_1)^2} - \frac{1}{(2m\omega + 2m_1\omega_1)^2} \right\}. \end{aligned}$$

To carry through the passage from $\wp(u)$ to $\sigma(u)$ in a rigorous manner would require the help of several theorems concerning series, the development of which, however, would take us outside the scope of this paper. We are content with having shown how to obtain the general expression for the function $\wp(u)$, given by the series (20), starting from Abel's fundamental ideas. This also gives the analytic expression for a general elliptic function. This gives the elements on which the complete theory of these functions can be founded.

The elements of the theory which we have presented are all to be found in Abel's memoir "*Recherches sur les fonctions elliptiques*", published in Crelle's Journal, vol. 2, no. 2, 1827, and vol. 3, no. 2, 1828.* The equalities (18), (19) and (20) in § 2, which express one of the fundamental ideas of Abel have not been given in such a general form by him. In "*Recherches* etc." only the

* Republished in Œuvres complètes, Holmboe's edition, vol. 1, page 141. [Page 263 in the second edition by Sylow and Lie. T. n.]

special case treated in formula (26), § 2 occurs. This is, of course, all that is needed for the special purpose on hand, but the true source of the imaginary transformation (26), § 2, being the more general formula (20), § 2, we prefer to start from the latter. The transformation (20), § 2, really belongs to a different field of the theory of elliptic functions. As a matter of fact, it contains, among other things, the theorem due to Jacobi to the effect that the multiplication problem for the elliptic functions can be solved by a two-fold application of a rational transformation.* The demonstration upon which Abel bases the passage from the multiplication theorem of the elliptic functions to the general analytic expression of these functions is, on one hand, exceedingly complicated, on the other hand, scarcely conclusive. We have tried to replace Abel's proof by another proof which is both simple and rigorous; moreover, we have not brought to our aid any other means than those which served Abel. Besides the formula analogous to (20) which Abel discovered, he also expressed the function $\wp(u)-e_1$, $\wp(u)-e_2$ and $\wp(u)-e_3$ (or rather functions analogous to these) as infinite products. This can be done either through the function $\sigma(u)$ or directly by a method similar to the one which gave (20). The way of handling the proof is evident from the treatment above.

We must not forget that the whole argument of Abel rests upon the assumption that the three roots of $S=0$ are *real*. This assumption can be replaced by the slightly more general assumption that *the two constants g_2 and g_3 are real*. In order to gain greater simplicity in the exposition we have restricted ourselves to the former assumption which means that in addition to g_2 and g_3 being real we suppose $g_2^3-27\,g_3^2>0$. It would not introduce any difficulties, however, to modify the argument in such a way that it holds also when $g_2^3-27\,g_3^2<0$.

The meaning of the function $\wp(u)$ and of the elliptic functions in general is, however, much wider, and $\wp(u)$ has still a definite meaning and preserves most of its essential properties when g_2 and g_3 are arbitrary complex quantities. The success of Abel's method of investigation, however, depends essentially upon the realness of g_2 and g_3, since the integral (15), § 2, has a definite meaning for us only under this assumption.† If we want to obtain a complete theory of the elliptic function following the road indicated by Abel, we have

* The formula (25) itself is, of course, due to Weierstrass.

† [In the second Abel memorial volume of Acta Mathematica, vol. 27 (1903), P. Mansion in a paper "Sur la méthode d'Abel pour l'inversion de la première integral elliptique, dans le cas où le module a une valeur imaginaire complexe", extends the method of Abel to the case when the coefficients of $R(x)$ are complex numbers. Complex integration in the ordinary sense of the word is avoided, but the author considers integrals of complex functions of a real variable. The method is applied to the Jacobian elliptic functions and not to the Weierstrassian ones. T. n.]

to call upon other resources. These are sufficiently well indicated by the character of the method. The natural continuation of our investigation would be to subject the general analytic expression obtained for $\wp(u)$ to a closer inspection. Such an inspection would show that this expression can be made to satisfy the differential equation (11), § 2, for arbitrary real or imaginary values of g_2 and g_3, provided the two periods 2ω and $2\omega_1$ are properly chosen. Thus the series (20) always gives the solution of equation (11) with the given initial condition. To develop the argument in detail would, however, carry us too far and we stop the investigation at this point. The elements thereof are sufficiently well indicated as soon as we can express $\wp(u)$ by means of the series (20).

At the side of Abel, Jacobi figures as one of the discoverers of the elliptic function as we have remarked in the introduction. The short notices in Crelle's and Schumacher's journals in which he announced his first discoveries, were followed in 1829 after Abel's death by the famous memoir *Fundamenta nova Theoriae Functionum Ellipticarum.** This paper contains a complete account of the most essential properties of the elliptic functions which had been obtained up to that time through the work of Abel and Jacobi. The abundance of new and ingenious ideas, the wealth of essential discoveries which are presented here on a few leaves, will always place this memoir among the classical works of mathematical literature.

However, we do not find a real mathematical theory of the elliptic functions in *Fundamenta* such as is given in Abel's words. To begin with, we miss altogether the kind of development which we have given in § 2 which shows that the function $\wp(u)$ exists for all values of the argument and, moreover, is a single-valued function of u. This, Jacobi considered to be self-evident. Considerable criticism could be directed against the line of thought which leads Jacobi to the general analytic expression of the elliptic functions. Abel considered himself obliged to investigate in detail the limiting process applied to the multiplication formula. Jacobi, on the other hand, carries through the limiting process without trying to justify the validity of the method. The possibility of carrying through the limiting process in Abel's theory depends upon the facts that $\wp(u)$ is a single-valued function of u and that the analytic expression of $\wp'(u)$ in terms of $\wp(u)$ is known. This gives an approximate analytic expression of $\wp(u)$ in the neighbourhood of $u = 0$, from which it is possible to pass over to the general expression. The limiting process in Jacobi's theory on the other hand is with respect to the constants g_2 and g_3. It is obvious that, at that stage of the theory before the general analytic expression of $\wp(u)$ has been obtained, it is impossible to determine whether $\wp(u)$ also is

* [Reprinted in Gesammelte Werke, vol. 1, pp. 49–239. T. n.]

a single-valued and analytic function of the parameters g_2 and g_3.* Thus it seems to be impossible at this stage of the development of the theory to undertake a closer investigation of such a limiting passage as that of Jacobi; at any rate if one desires a Jacobian theory which is essentially distinct from that of Abel. It is possible to carry through the limiting process with the aid of Abel's multiplication formulae but then the theory of Jacobi would differ from that of Abel only by a higher degree of complication. A closer analysis of *Fundamenta novae* from the point of view which we have tried to indicate would be of great interest and we intend to return to this problem at some other occasion. Such an investigation would be out of place here.

The function $\wp(u)$ which is one of the most remarkable functions known to mathematics, was introduced by Weierstrass.† In his *Recherches*, — etc., Abel uses three different functions $\varphi(u)$, $f(u)$ and $F(u)$ instead of $\wp(u)$. In his *Fundamenta novae* Jacobi considers three other functions sin am u, cos am u

* [This problem can be attacked by the methods of Poincaré for the study of the dependence of the solutions of differential equations upon parameters entering in the equation. Cf. Poincaré: Les méthodes nouvelles de la mécanique céleste, vol. 1, p. 51 et seq. Also Picard: Traité d'Analyse, vol. 3, Chap. 8. T. n.]

† [The only presentation of Weierstrass' theory of the $\wp$-function which appeared during his life-time was the outline "Formeln und Lehrsätze zum Gebrauche der Elliptischen Functionen", edited by H. A. Schwarz, Berlin (1892). The fifth volume of his Mathematische Werke, "Vorlesungen über die Theorie der Elliptischen Funktionen", (1915), edited by J. Knoblauch, gives an exposition of the theory based on Weierstrass' lectures during various years. As Weierstrass repeatedly lectured upon the theory of elliptic functions during his long career as lecturer and naturally slightly varied his selection of material from year to year, even this admirable presentation does not give a complete account of the theory. So for instance, we greatly miss the systematic treatment of the addition theorem from the point of view of a functional equation defining a class of analytic functions, with which Weierstrass often started his exposition.

Any modern treatise on analysis or function theory will contain a chapter devoted to the $\wp$-function. Of the many monographs written on the theory of elliptic functions the more recent ones will give a more or less complete account of the essential properties of the $\wp$-function. A few such monographs are listed below.

Appell et Lacour: Principes de la Théorie des Fonctions Elliptiques, Paris (1897).

Burckhardt: Functionentheoretische Vorlesungen, Zweiter Teil, Elliptische Functionen, Leipzig (1899).

Fricke: Die elliptischen Funktionen und ihre Anwendungen, Leipzig (1916 and 1922, two volumes).

Hancock: Lectures on the theory of Elliptic Functions, New York (1910).

Tannery et Molk: Éléments de la Théorie des Fonctions Elliptiques, Paris (1893, 1896, 1898 and 1902, four volumes).

Weber: Elliptische Funktionen und Algebraische Zahlen (vol. 3 of Lehrbuch der Algebra) Braunschweig (1908).

All these books also present the theory of Jacobi's functions and are perhaps permeated more by the spirits of Cauchy and Riemann than that of Weierstrass. T. n.]

and Δ ampl u. These functions with the notation of Jacobi* have been incorporated with mathematical literature. We have let the theory of Abel lead up to the function $\wp(u)$ instead of to the three functions of Abel or Jacobi. This is no essential change but gives the investigation a considerable formal simplification. Once in possession of the function $\wp(u)$, it is an easy matter to pass over to the functions of Abel and Jacobi. Another reason which has been decisive in our choice of ways and means in dealing with the problem, is to get our work to connect up with the profound and highly developed theory of the elliptic functions which is due to Weierstrass. This theory, as the mathematical world is well aware, is centered around the function $\wp(u)$ and the function $\sigma(u)$ derived from the former. If this theory is supposed to be known, the most natural course in presenting Abel's line of thought is to use $\wp(u)$ instead of the functions of Jacobi.

§ 4.

The deduction, according to Weierstrass, of the general solution of the differential equation

$$\frac{dx}{\sqrt{R(x)}} = du, \tag{1}$$

$$R(x) = Ax^4 + 4Bx^3 + 6Cx^2 + 4B'x + A'. \tag{2}$$

With an ingenuity which is hard to match in the history of mathematics, Abel had succeeded in finding the elementary principles upon which a rigorous theory of elliptic functions could be founded. While the theory of Abel solves the problem of deducing and analyzing the elliptic functions with elementary means in a simple and thorough manner, it suffers from certain imperfections which are consequences of just the elementary nature of the method. Such an imperfection is the fact that *two* different principles are necessary; besides the addition theorem the imaginary transformation of (26), § 2, has to support the argument. Another imperfection lies in the failure of the method to give us directly the elliptic functions for complex values of g_2 and g_3. Still there is no essential difference between the functions which correspond to *real* and to *complex* values of the constants.

It became possible to remedy these deficiencies and to give Abel's theory a higher degree of perfection when Weierstrass showed† how to determine

* [Nowadays the notation of Jacobi is frequently replaced by the abbreviated notation of Gudermann, viz. sn u, cn u and dn u for the functions mentioned above. T. n.]

† "Über die Theorie der analytischen Facultäten", Journal f. Mathematik, vol. 51, p. 1 et seq. (1854). The theorem is stated without proof on pp. 43–44. This paper was later

the general solution of a system of simultaneous differential equations. It is true that by adjoining this idea Abel's method loses its elementary character and will be founded to a considerable extent upon certain general theorems of function theory. On the other hand, what the method loses in simplicity, it will gain in unity and lucidity. The outline of the theory was presented by Weierstrass in *Theorie der Abelschen Functionen*. Journal f. Mathematik, vol. 52, p. 62 et seq. (1856).

Instead of equation (1) we may introduce the system of simultaneous equations

$$(3)\qquad \frac{dx}{du} = x', \qquad \frac{dx'}{du} = \tfrac{1}{2} R'(x).$$

From the existence theorem of Weierstrass, it follows that in the immediate neighbourhood of $u = 0$ we can form the general solution of these equations. This solution is given by the power series

$$(4)\qquad x = a + \sum_{n=1}^{\infty} \left(\frac{d^n x}{du^n}\right)_{u=0} \frac{u^n}{n!}$$

and the derived series for x'. Here a denotes an arbitrary constant. It is always possible to find a finite positive quantity ϱ such that the series (4) is *absolutely convergent* within the region

$$(5)\qquad |u| < \varrho.$$

Here the quantity ϱ must be chosen so small that neither of the variables x and x' nor their first derivatives becomes infinite within the region determined

reprinted, first in Abhandlungen aus der Funktionenlehre (1886), secondly in Mathematische Werke, vol. 1 (1894). In both these later editions the theorem was left out. Volume 1 of Werke contains a paper "Definition Analytischer Functionen einer Veränderlichen vermittelst Algebraischer Differentialgleichungen", pp. 75–85, written in 1842 but up to that time unpublished. This is evidently the investigation the result of which Weierstrass quotes in his paper of 1854. The proof follows without difficulty from the other results of the latter paper, and was probably known in wide circles in Germany through private communication. Cf. the thesis of Mme. Sophie von Kovalewsky, "Zur Theorie der partiellen Differentialgleichungen", Journal f. Mathematik, vol. 80, pp. 1–32 (1875).

The first published proof of the theorem was given by Briot and Bouquet, "Recherches sur les propriétés des fonctions définies par des équations differentielles", Journal de l'École polytechniques, cap. 36, pp. 133–198 (1856). [Note revised by translator.]

by (5). But neither x' nor $\dfrac{dx'}{du}$ can become infinite except when x is infinite. Thus the number ϱ is the *radius of convergence* of the power series (4). From equation (1) it follows immediately that

$$(6)\qquad \frac{d^{2n-1}x}{du^{2n-1}} = R_{2n-1}(x)\,\frac{dx}{du}.$$

$$(7)\qquad \frac{d^{2n}x}{du^{2n}} = R_{2n}(x).$$

where n is a positive integer and R_{2n-1} and R_{2n} are *integral* rational functions of x. Thus the series (4) can be written:

$$(8)\qquad x = a + \sum_{n=1}^{\infty} R_{2n}(a)\,\frac{u^{2n}}{(2n)!} + \sqrt{R(a)}\,\sum_{n=1}^{\infty} R_{2n-1}(a)\,\frac{u^{2n-1}}{(2n-1)!}.$$

In view of § 1. we can transform the differential equation (1) into the equation

$$(9)\qquad -\frac{ds}{\sqrt{S}} = du, \qquad S = 4s^3 - g_2 s - g_3.$$

If we introduce the substitution

$$(10)\qquad s = \frac{1}{z},$$

we get

$$(11)\qquad ds = -\frac{dz}{z^2}.$$

Hence equation (9) becomes

$$(12)\qquad \left(\frac{dz}{du}\right)^2 = z(4 - g_2 z^2 - g_3 z^3).$$

If we integrate the equation (12) with the initial condition

$$u = 0, \quad z = 0,$$

we obtain the solution of (9) with the initial condition

$$u = 0, \quad s = \infty.$$

Observing that the value of $R(a)$ corresponding to equation (12) is zero, (8) yields

$$(13) \qquad z = \sum_{n=1}^{\infty} R_{2n}(a) \frac{u^{2n}}{(2n)!} = u^2 - \frac{g_2}{4 \cdot 5} u^6 - \frac{g_3}{4 \cdot 7} u^8 - \cdots.$$

Thus the general solution of equation (9) is given by

$$(14) \qquad s = \wp(u) = 1 \Big/ \left\{ u^2 - \frac{g_2}{4 \cdot 5} u^6 - \frac{g_3}{4 \cdot 7} u^8 - \cdots \right\}.$$

If we let δ denote the smallest of the absolute values of the roots of the series (13) equated to zero, we have, for all values of u which are less than δ in absolute value

$$(15) \qquad s = \wp(u) = \frac{1}{u^2} + \frac{g_2}{4 \cdot 5} u^2 + \frac{g_3}{4 \cdot 7} u^4 + \cdots,$$

where *the power series is absolutely convergent when* $|u| < \delta$ *and this number* δ *is the precise radius of convergence of the series.*

Let us integrate twice, with appropriate choice of the constants of integration, forming the series

$$(16) \quad G(u) = \int_0^u \int_0^v \left[\wp(t) - \frac{1}{t^2} \right] dv\, dt = \frac{g_2}{2^4 \cdot 3 \cdot 5} u^4 + \frac{g_3}{2^3 \cdot 3 \cdot 5 \cdot 7} u^6 + \cdots,$$

and

$$(17) \qquad \begin{aligned} \sigma(u) &= u e^{-G(u)} = u - \frac{g_2}{2^4 \cdot 3 \cdot 5} u^5 - \frac{g_3}{2^3 \cdot 3 \cdot 5 \cdot 7} u^7 - \cdots \\ &= u \left\{ 1 - \frac{g_2}{2} \frac{u^4}{5!} - 6 g_3 \frac{u^6}{7!} - \cdots \right\}. \end{aligned}$$

This series is also absolutely convergent, at least when u is within the circle $|u| = \delta$. It is uncertain whether the radius of convergence of the power series is δ or is greater than δ.

Abel starts with an element of $\wp(u)$ obtained from the elliptic integral, out of this element the function is formed by means of the imaginary transformation and the addition theorem. Weierstrass, on the other hand, starts from the element (15) which has been obtained directly from the general theory of differential equations, and from this element he forms the function by using the addition theorem in a manner similar to Abel's.

If we consider the addition formula (44) § 2, and put $u = v$ the formula takes on an indeterminate form which is easily reduced to

$$(18) \qquad 4\wp(2u) = 4\wp(u) - \frac{d}{du}\left\{\frac{\wp''(u)}{\wp'(u)}\right\}.$$

Instead of investigating this formula directly we integrate twice. Observing that

$$(19) \qquad -\frac{d^2 \log \sigma(u)}{du^2} = \wp(u)$$

and choosing the correct constants of integration, we obtain a multiplication formula for the function $\sigma(u)$, namely

$$\sigma(2u) = -[\sigma(u)]^4\, \wp'(u) = [\sigma(u)]^4 \cdot \frac{d^3 \log \sigma(u)}{du^3}.$$

Replacing $2u$ by u we get

$$(20) \qquad \sigma(u) = -8\left[\sigma\left(\frac{u}{2}\right)\right]^4 \cdot \frac{d^3}{du^3}\left[\log \sigma\left(\frac{u}{2}\right)\right].$$

Carrying through the indicated differentiation, we find

$$(21) \quad \sigma(u) = -\sigma'''\left(\frac{u}{2}\right)\sigma^3\left(\frac{u}{2}\right) - 3\sigma''\left(\frac{u}{2}\right)\sigma'\left(\frac{u}{2}\right)\sigma^2\left(\frac{u}{2}\right) + 2\left[\sigma'\left(\frac{u}{2}\right)\right]^3\sigma\left(\frac{u}{2}\right).$$

Thus the functin $\sigma(u)$ is an integral rational function of $\sigma\left(\frac{u}{2}\right)$ and of the first three derivatives of that function.

The function $\sigma(u)$ is represented by an absolutely convergent power series at least within the circle $|u| = \delta$. Hence the function $\sigma\left(\frac{u}{2}\right)$ is represented by such a power series at least within the circle $|u| = 2\delta$. The derivatives of $\sigma\left(\frac{u}{2}\right)$ are likewise represented by absolutely convergent power series in u

at least within the circle $|u| = 2\delta$. The right hand side of equation (21) is formed by sums and products of a finite number of power series, all of which converge at least in $|u| = 2\delta$. Thus the right hand side itself is expressible as an absolutely convergent power series within $|u| = 2\delta$. About the left hand side of (21) we know with certainty only that it is a power series convergent when $|u| < \delta$. If two absolutely convergent power series in u are equal within a region common to their circles of convergence, they are *identically* equal and since they proceed according to powers of the same variable, they have the same region of convergence. Thus the power series for $\sigma(u)$ about which we originally know only that it was convergent when $|u| < \delta$, we have now shown to be convergent within the greater region $|u| = 2\delta$ on account of relation (21). Repeating the argument we conclude that the power series (17) for $\sigma(u)$ is absolutely convergent for every finite value of u, or in other words $\sigma(u)$ *is an everywhere convergent power series.**

This result is very remarkable. We point out only one circumstance. The power series for $\sigma(u)$ we obtained by forming $u\ exp\,[-G(u)]$, or in other words by replacing the independent variable in the exponential series by the series $-G(u)$ and multiplying the result by u. *But the series for $G(u)$ is divergent outside of the region $|u| = \delta$, still the result of the substitution when rearranged is an everywhere convergent power series in u.* The method through which we found that $\sigma(u)$ is an everywhere convergent power series is exceedingly fruitful. It is capable of considerable generalization and has lead Weierstrass to certain theorems which belong to the deepest and most comprehensive results of present day mathematics.†

Once we have got the function $\sigma(u)$ defined for every finite value of u, the function $\wp(u)$ is also completely determined. As a matter of fact, it can be shown that formula (19) which originally held only when $|u| < \delta$, remains true for every finite value of u. The functions $\wp(u)$ and $\sigma(u)$ which we obtain directly by Weierstrass' method are the general functions *Pe* and *Sigma* and *the corresponding constants g_2 and g_3 are arbitrary real or complex quantities.*

It is true that formula (17) tells us that $\sigma(u)$ is developable in a power series which we later proved to be everywhere convergent. The formula, however, does not give us any information about the law of formation of the

* In our presentation we have simplified the proof which Weierstrass gives in "Theorie der Abelschen Funktionen". In order to proceed according to Weierstrass we should have developed the formula for $\sigma(nu)$ and then drawn the same conclusions which we have based on the simpler formula for $\sigma(2u)$. The simplified treatment which we have given is developed on account of a verbal message from Weierstrass that he, after the publication of "Theorie der Abelschen Funktionen", had simplified his method by replacing the latter n throughout by the number 2. In his original paper Weierstrass used the functions $Al(u)$ instead of the later discovered functions $\wp(u)$ and $\sigma(u)$.

† [The author has in mind the factorization theorem for entire functions. T. n.]

coefficients which, by the way, is rather difficult to obtain; thus we can not be said to be in real possession of the function $\sigma(u)$. Weierstrass showed in the paper *Theorie der Abelschen Funktionen*, repeatedly quoted, how the law of formation of the coefficients could be obtained, thus enabling us to actually write down the power series of $\sigma(u)$. In the same memoir Weierstrass gave a second method which leads to another analytic expression for $\sigma(u)$ and thence also for $\wp(u)$. As a matter of fact, given that $\wp(u)$ is doubly periodic and the corresponding properties of $\sigma(u)$ which follow from the addition theorem according to Abel, Weierstrass showed how to express $\sigma(u)$ and $\wp(u)$ in terms of the Jacobian Θ-functions. To these two methods we can now add a third one, since we have shown that Abel's method of expressing the $\wp$-function by the series (20), § 3, through a limiting process applied to the multiplication theorem, can be carried through rigorously. It is obvious that this procedure is applicable whether we start from the basis furnished by Weierstrass' method set forth in this paragraph or we continue Abel's investigation presented in § 2.

The historical source of the theory of elliptic functions was a perfectly definite restricted problem, namely to solve the differential equation

$$\frac{dx}{\sqrt{Ax^4+4Bx^3+6Cx^2+4B'x+A'}} = du. \tag{22}$$

This problem has a perfectly general solution, thus the equation (22) contains the definition of the class of functions which have received the name *elliptic functions*.

Jacobi himself was conscious of the fact that the method he used in *Fundamenta nova* was unsatisfactory, and he very soon abandoned this method and used an entirely different starting point for the theory of elliptic functions in his lectures. He started from the study of some of the general analytic expressions for the elliptic functions. This study led to the functions themselves and their more important properties. Unfortunately Jacobi's lectures have not been completely published* but their essential content is fairly well known through communications by himself and by his pupils, also by the lectures of the latter and by copies of notes. A method similar to the later method of Jacobi was used by Hermite† as early as 1862, who gave a masterly

* [In Jacobi's Gesammelte Werke, vol. 7, page 412, appears a list of notes from Jacobi's lectures which are in the possession of the Academy of Sciences at Berlin. Five of these sets of notes are concerned with elliptic functions. T. n.]

† Cf. for this matter two letters from Hermite to Jacobi and the answer of Jacobi in Jacobi's Gesammelte Werke, vol. 2, pp. 87–120. These letters were first published in Journal f. Mathematik, vol. 32, pp. 176–181, 277–299 (1845).

compendium of his method of deducing the elliptic functions in *Note sur la théorie des fonctions elliptiques*, published in the second volume of the sixth edition of Lacroix: *Traité élémentaire de calcul différentiel et de calcul intégral.* To these methods of Jacobi and Hermite the objection might be made that the starting point of the theory is rather arbitrary and it is difficult to find a unique point of view on the study of the theory of elliptic functions when approached in that way.* Jacobi was perfectly aware of this deficiency; in his lectures he was accustomed to point out the advantage in choosing the differential equation (22) as starting point of the theory, and remarked that he intended to return to that basis as soon as the theory of complex integration had advanced so far that a rigorous theory could be built on that foundation.†

The theory of complex integration was brought to perfection, about the need of which Jacobi held such strong beliefs, through the epochal work of Cauchy. An abstract of this theory together with further developments was published by Puiseux in Journal de Mathématiques, vol. 15–16, (1850–51). Briot and Bouquet based their treatise *Théorie des fonctions doublement périodiques et, en particulier, des fonctions elliptiques*, (1859) on Cauchy's theories.‡

However, and this is well worth noticing, even the first paper by Abel contained a theory of elliptic functions, which, save for various unsatisfactory details, fulfils all the requirements that in reason may be put on a mathematical theory. We have seen what further perfection can be given to Abel's theory through the methods of Weierstrass. Still there is one objection that can be raised against any theory of elliptic functions based upon the differential equation (22). Modern mathematics as it has been formed by the hands of the great masters is rightly exacting in its requirements on definitions. The definition of a new class of functions ought to be so chosen that it contains no restriction the real explanation of which is not possible to give a priori. The differential equation (22) regarded as a definition of the elliptic functions does not fulfil this requirement. Why should we restrict the polynomial under the radical to be of degree four at most, and why just a square-root and no other algebraic function? The real answers to these questions are only given in the far more advanced theory of *Abelian functions*.

* Perhaps such a unifying point of view could be obtained in the arithmetic-geometric mean of Gauss. The study of this expression seems to have led Gauss to his discovery of the elliptic functions.

† We owe this information to a remark by Weierstrass.

‡ The investigations of Liouville with which we shall be concerned below, also play a fundamental rôle in this treatise. A second edition of this book appeared in 1875 under the title Théorie des fonctions elliptiques, in which the authors for reasons unknown have largely abandoned the elegant and methodical theory that they adherred to in the first edition.

In the epoch-making memoir *De functionibus duarum variabilium quadrupliciter periodicis, quibus theoria transcendentium Abelianarum innititur,** Jacobi showed that *a function of a single variable which for every finite value of the argument has the character of a rational function*† *cannot have more than two periods and the ratio of the periods is necessarily a non-real number.* This discovery permitted a new point of view on the theory of the elliptic functions, namely the problem of finding all *doubly periodic functions.* As early as in 1844 Liouville in a communication to the French Academy of Sciences‡ had shown how to develop a complete theory of *doubly periodic functions* from this starting point. Right at the start of the theory a whole set of new general theorems were obtained. This discovery of the great geometer was the first contribution of fundamental importance to the theory of elliptic functions since their introduction in mathematics by Abel and Jacobi. In the *double periodicity* Liouville had discovered an essential property of the elliptic functions and a unifying point of view to apply to their theory. Moreover, this starting point is not subject to the kind of objection which we raised against the choice of equation (22). The *doubly periodic* functions form a more general class of functions than those designated by Jacobi by the name *elliptic.* This is also a great advantage, especially as all the fundamental properties of *elliptic* functions recur with the *doubly periodic* functions. The particular determination through which the elliptic functions are singled out from the doubly periodic functions is of rather incidental nature.

One remark may still be made against the choice of the *double periodicity* as the starting point for the theory of these functions. The nearest lower class of functions would then be the *simply periodic* functions. These, however, are a very much more general class of functions, and it is only a small sub-class of these functions which can be obtained by specializations of the double periodicity. Then the question arises: It is perhaps possible to find a characteristic property which is common to the doubly periodic functions and this particular sub-class of simply periodic functions, and which is the exclusive property of these functions and thus distinguishes them in contract to all other analytic functions? Weierstrass found such a property in the *addition theorem.*

* Journal f. Mathematik, vol. 13, pp. 55–78 (1835). [Reprinted in Gesammelte Werke, vol. 2, pp. 23–50, T. n.]

† [*I. e.* a meromorphic function. T. n.]

‡ C. R., vol. 19, p. 1261, See further C. R., vol. 32, "Rapport sur un mémoire à l'Académie par M. Hermite et relatif aux fonctions a double période (Commissaires M. M. Sturm, Cauchy rapporteur) and Remarques de M. Liouville and Note de M. Augustin Cauchy relative aux observations présentés à l'académie par M. Liouville." It is peculiar that Liouville never published a detailed account of these discoveries which belong to the most important of his many contributions to mathematics.

and from this starting point he succeeded in developing the theory of elliptic functions to the highest degree of perfection that a mathematical theory may ever reach.* Weierstrass starts out with the following general problem which is sufficiently well justified a priori by its importance in the elementary theory of *algebraic* and *trigonometric* functions, namely:

An element of an analytic function, given within a certain convex region, has the following property. There exists an algebraic relation between three values of the element, of which two correspond to arbitrary values of the argument within the given region and the third corresponds to the arithmetic mean of these two arguments. The coefficients of this relation do not depend upon the choice of the points in the region. Which functional element has this general characteristic property and to what function does it belong?

The answer to this question is:

The element will have this characteristic property if and only if it is the root of an algebraic equation the coefficients of which are rational functions of $\wp(u)$ *and* $\wp'(u)$.

---

* This second method of Weierstrass is entirely different from that used in "Theorie der Abelschen Funktionen". [This method has been made known chiefly through the publications of the pupils of Weierstrass. There does not yet seem to exist a systematic treatise on elliptic functions based entirely on the addition theorem. The latter theorem, however, plays a fundamental rôle in Hancock's Theory of Elliptic Functions. Various proofs of the fundamental theorem have been pubished. Cf. Phragmén, "Sur un théorème concernant les fonctions elliptiques", Acta Mathematica, vol. 7, pp. 33–42, (1885). Further the thesis of P. Koebe, "Über diejenigen analytischen Funktionen eines Arguments, welche ein Algebraisches Additionstheorem besitzen" (Berlin, 1905). Reprinted in extended and revised form in Mathematische Abhandlungen Hermann Amandus Schwarz, etc. gewidmet, pp. 192–214 (Berlin, 1914). Also M. Falk, "Über die Haupteigenschaften derjenigen Funktionen eines Arguments, welche Additionstheoreme besitzen", Nova Acta, soc. Upsal., ser. 4, vol. 1, no. 8 (1907). T. n.]

CYCLOTOMIC QUINQUISECTION FOR ALL PRIMES OF THE FORM $10n+1$ BETWEEN 1900 AND 2100.

BY PANDIT OUDH UPADHYAYA.*

Legendre† considered the problem of cyclotomic quinquisection for the prime 641 and calculated the corresponding period equation‡

$$\eta^5 + \eta^4 - 256\eta^3 - 564\eta^2 + 5328\eta - 5120 = 0.$$

The problem of cyclotomic quinquisection was attempted by Cayley§ in two papers in the Proceedings of the London Mathematical Society but he was unable to solve the problem completely. The same problem was also considered by Lloyd Tanner. Miss C. A. Scott¶ considered the problem of quarti-section and quinquisection but she did not succeed in finding the expressions for the constant term in the period equation. The problem of quinquisection was completely solved by Rogers** who showed that it depends on the solution of two diophantine equations. The same problem has been considered recently by Burnside.†† He showed that it depends on the solution of the diophantine equations

$$(1)\quad 12^2 p = [4p - 16 - 25(A+B)]^2 + 1125(A-B)^2 + 450(C^2 + D^2),$$

$$(2)\quad 0 = [4p - 16 - 25(A+B)]\,[A-B] + 3(C^2 + 4CD - D^2),$$

and he gives the period equation

---

* Babu Shiva Prasad Gupta Research Scholar.

† Theorie des nombres, 3e ed., t. 2, p. 213.

‡ The misprint in the coefficient of η has been corrected.

§ "The binomial equation $x^p - 1 = 0$: Quinquisection", Proc. Lond. Math. Soc., vol. 12 (1880), pp. 15, 16, and vol. 16 (1885), pp. 61-63.

‖ Proc. Lond. Math. Soc., vol. 18 (1887) pp. 214-234.

¶ Amer. Jour. Math., vol. 8 (1886), pp. 261-264.

** Proc. Lond. Math. Soc., vol. 32 (1900), pp. 199-207.

†† Proc. Lond. Math. Soc., Ser. 2, vol. 14, pp. 251-259.

$$
\begin{aligned}
&\eta^5+\eta^4-\frac{2}{5}(p-1)\eta^3+\left[\frac{1}{3}p(A+B)-\frac{2(p-1)(2p+3)}{3\cdot 5^2}\right]\eta^2\\
&\quad+\left[\frac{p}{9}\left(\frac{p-1}{5}+A+B\right)^2-pAB-\frac{(p-1)^3}{5^3}\right]\eta\\
(3)\qquad&\quad+\frac{p}{5}\left[\frac{1}{5\cdot 6^3}\left\{5(A+B)-\frac{4p-4}{5}\right\}^3+\frac{1}{6^2}\left\{\frac{2p-2}{5}-A-B\right\}^2\right.\\
&\quad\left.+\frac{1}{4}(A-B)^2+\frac{1}{8}(A-B)(D^2-C^2)\right]-\frac{(p-1)^3}{5^5}=0.
\end{aligned}
$$

These three equations were first given by Rogers in a different notation but in this paper I have adhered to the notation of Burnside. So far as I am aware the first two equations have been solved for the primes 11, 31, 41, 61, and 71 but the period equation has not been given even for all of these numbers. The object of this paper is to solve the two diophantine equations for all primes of the form $10n+1$ between 1900 and 2100 and to calculate the corresponding period equations.

The prime 1901. If in the first equation we substitute the value of p, supposing that $A+B=307$, we get

$$[4\times 1901-16-25\times 307]^2+1125(A-B)^2+450(C^2+D^2)=144\times 1901,$$

or

$$1125(A-B)^2+450(C^2+D^2)=273744-7569.$$

If $A-B=11$, this gives

$$C^2+D^2=289=8^2+15^2.$$

We therefore have

$$A=159,\quad B=148,\quad C=8,\quad D=15.$$

and on substitution it is found that these values satisfy the second equation. Substituting in (3), we find the period equation to be.

$$\eta^5+\eta^4-760\eta^3+1749\eta^2+84009\eta+69211=0.$$

The prime 1931. Solving the first two equations we find

$$A = 159, \quad B = 160, \quad C = 4, \quad D = 21.$$

The period equation is

$$\eta^5 + \eta^4 - 772\,\eta^3 + 6411\,\eta^2 + 2379\,\eta - 72047 = 0.$$

The prime 1951. Solving the first two equations, we find

$$A = 160, \quad B = 155, \quad C = 23, \quad D = -4.$$

The period equation is

$$\eta^5 + \eta^4 - 780\,\eta^3 + 1795\,\eta^2 + 40175\,\eta - 150571 = 0.$$

The prime 2011.

$$A = 158, \quad B = 154, \quad C = 2, \quad D = 22.$$

The periodic equation is

$$\eta^5 + \eta^4 - 804\,\eta^3 - 6596\,\eta^2 + 14624\,\eta + 162752 = 0.$$

The prime 2081.

$$A = 165, \quad B = 166, \quad C = 6, \quad D = 25.$$

The period equation is

$$\eta^5 + \eta^4 - 832\,\eta^3 - 1415\,\eta^2 + 34195\,\eta - 8621 = 0.$$

I should like to mention that I have received help in calculation from Pandit Shukdeo Chanbey.

GEODESIC LINES IN RIEMANN SPACE.

By R. Henderson.

When a, b and c are the sides of a triangle and C the angle opposite to c, we have

$$c^2 = a^2 + b^2 - 2ab\cos C.$$

Differentiating with respect to b, we have

$$2c\frac{dc}{db} = 2b - 2a\cos C = 2c\cos A,$$

$$\frac{dc}{db} = \cos A.$$

Similarly in a spherical triangle

$$\cos c = \cos a\cos b + \sin a\sin b\cos C,$$

$$\sin c\frac{dc}{db} = \cos a\sin b - \sin a\cos b\cos C$$

$$= \sin c\cos A,$$

$$\frac{dc}{db} = \cos A.$$

The differential coefficient is therefore in each case independent of the magnitude of c provided the angle A is given. The sides of the triangle are in each case geodesic lines. Therefore, this gives us a property of geodesic lines which for the purposes of this investigation it is proposed to take as fundamental. It may be expressed as follows: If P and P' be any two points and s be the distance between them, then the variation of s, due to any arbitrary infinitesimal change in the position of P, is independent of the actual magnitude of s and depends only on the direction components at P of the geodesic line joining the two points. This property appears to be closely associated with the fact that the geodesic is the line along which the distance between any two points on it is measured so that at any point the gradient of the distance is independent both in magnitude and direction of the starting point *on the line* from which it is measured. We shall designate coordinates and functions at the point P by unaccented letters and those at P' by accented letters.

In ordinary cartesian coordinates for space of three dimensions we have

$$s^2 = (x-x_1)^2+(y-y_1)^2+(z-z_1)^2.$$

Then if dl represent the length of the infinitesimal displacement of P, we have

$$\begin{aligned}
2s\frac{ds}{dl} &= 2(x-x_1)\frac{dx}{dl}+2(y-y_1)\frac{dy}{dl}+2(z-z_1)\frac{dz}{dl},\\
\frac{ds}{dl} &= \frac{x-x_1}{s}\cdot\frac{dx}{dl}+\frac{y-y_1}{s}\cdot\frac{dy}{dl}+\frac{(z-z_1)}{s}\cdot\frac{dz}{dl},\\
&= u_1\frac{dx}{dl}+u_2\frac{dy}{dl}+u_3\frac{dz}{dl},
\end{aligned}$$

where u_1, u_2 and u_3 are the direction cosines of the straight line drawn from P' to P. They are, of course, independent of the length s and depend only on the particular straight line through P on which P' is situated. The value of ds/dl is therefore independent of s. Since dx/dl, dy/dl and dz/dl are the direction cosines of the displacement of P, the expression obtained for ds/dl is evidently equal to the cosine of the angle which the direction of displacement of P makes with the line from P' to P.

In the case of Euclidean space of n dimensions, referred to oblique rectilinear coordinates x^i, x^k etc., we have

$$s^2 = \sum_{ik} a_{ik}(x^i-x_1^i)(x^k-x_1^k), \qquad (a_{ik} = a_{ki}),$$

where the coefficients a_{ik} are constants. Then we have generally

$$\frac{ds}{dl} = \sum_r \frac{\partial s}{\partial x^r}\cdot\frac{dx^r}{dl}.$$

Our fundamental property therefore means that values of $\partial s/\partial x^r$ are independent of the magnitude of s, for all values of r.

$$\begin{aligned}
2s\frac{\partial s}{\partial x^r} &= \sum_i a_{ir}(x^i-x_1^i)+\sum_k a_{kr}(x^k-x_1^k),\\
&= 2\sum_i a_{ir}(x^i-x_1^i),\\
\frac{\partial s}{\partial x^r} &= \sum_i a_{ir}\frac{x^i-x_1^i}{s}.
\end{aligned}$$

Here again, if P' lies on a fixed straight line through P, $(x^i-x_1^i)/s$, which we shall call u^i, is a constant for all positions of P' on that line. Consequently, $\partial s/\partial x^r$ is also a constant.

Conversely, if the values of $\partial s/\partial x^r$ and consequently those of $(x^i - x_1^i)/s = u^i$ are constants, we have for each value of i

$$x_1^i = x^i - su^i.$$

Thus the coordinates of P' are linear functions of a single variable s and therefore the locus of P' must be a straight line.

Taking up now the general case of a Riemann space of n dimensions, the expression for an infinitesimal element of distance is

$$ds^2 = \sum_{ik} g_{ik}\, dx^i\, dx^k,$$

where the g_{ik} coefficients are not now constants but are functions of the coordinates. Here we cannot apply the expression directly to finite distances but for small distances δs we may write

$$\delta s^2 = \sum_{ik} \tfrac{1}{2}(g_{ik} + g'_{ik})\, \delta x^i\, \delta x^k,$$

where, generally, $\delta x^i = x^i - x_1^i$. It will be noted that the neglected difference between the two sides of this equation is of the order δs^4.

Then

$$2\,\delta s \frac{\partial(\delta s)}{\partial x^r} = \sum_i (g_{ir} + g'_{ir})\,\delta x^i + \tfrac{1}{2}\sum_{ik} \frac{\partial g_{ik}}{\partial x^r}\,\delta x^i\,\delta x^k,$$

$$\frac{\partial(\delta s)}{\partial x^r} = \sum_i \tfrac{1}{2}(g_{ir} + g'_{ir})\frac{\delta x^i}{\delta s} + \frac{\delta s}{4}\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}\cdot\frac{\delta x^i}{\delta s}\cdot\frac{\delta x^k}{\delta s}$$

$$= \sum_i g_{ir}\frac{dx^i}{ds} - \frac{\delta s}{2}\left\{\frac{d}{ds}\sum_i g_{ir}\frac{dx^i}{ds} - \tfrac{1}{2}\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}\cdot\frac{dx^i}{ds}\cdot\frac{dx^k}{ds}\right\},$$

if we expand in powers of δs and neglect powers above the first. This is independent of the magnitude of δs if

$$\frac{d}{ds}\sum_i g_{ir}\frac{dx^i}{ds} = \tfrac{1}{2}\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}\cdot\frac{dx^i}{ds}\cdot\frac{dx^k}{ds}.$$

If we write u^i for dx^i/ds, where u^i is now not necessarily constant, and u_r for $\sum_i g_{ir}\,u^i$ or $\sum_i g_{ir}\,dx^i/ds$, this becomes

$$\frac{du_r}{ds} = \tfrac{1}{2}\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}u^i\,u^k.$$

Another geometrical interpretation of the equation

$$\frac{d}{ds}\sum_i g_{ir}\frac{dx^i}{ds} = \tfrac{1}{2}\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}\cdot\frac{dx^i}{ds}\cdot\frac{dx^k}{ds}$$

may be obtained by integrating with respect to s, when we have

$$\sum_i g_{ir}\frac{dx^i}{ds}-\sum_i g'_{ir}\frac{dx^i_1}{ds}=\int_{P'}^{P}\tfrac{1}{2}\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}\cdot\frac{dx^i}{ds}\cdot\frac{dx^k}{ds}\,ds.$$

The coordinates of a variable point on the line and the distance along the line from an arbitrary origin to the point may be considered as functions of a single parameter l. Suppose now that the position of the line is varied by making at every point the same infinitesimal change ϵ^r in the x^r coordinate and leaving the other coordinates unchanged. Then, since

$$\left(\frac{ds}{dl}\right)^2=\sum_{ik}g_{ik}\frac{dx^i}{dl}\cdot\frac{dx^k}{dl}$$

and the values of dx^i/dl are independent of ϵ^r, we have

$$2\frac{ds}{dl}\,\frac{d^2s}{d\epsilon^r dl}=\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}\cdot\frac{dx^i}{dl}\cdot\frac{dx^k}{dl},$$

$$\frac{d^2s}{d\epsilon^r dl}=\tfrac{1}{2}\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}\cdot\frac{dx^i}{ds}\cdot\frac{dx^k}{ds}\cdot\frac{ds}{dl},$$

$$\int_{P'}^{P}\tfrac{1}{2}\sum_{ik}\frac{\partial g_{ik}}{\partial x^r}\cdot\frac{dx^i}{ds}\cdot\frac{dx^k}{ds}\,ds=\int_{P'}^{P}\frac{d^2s}{d\epsilon^r dl}\,dl=\frac{d}{d\epsilon^r}\int_{P'}^{P}ds.$$

Hence

$$\frac{d}{d\epsilon^r}\int_{P'}^{P}ds=\sum_i g_{ir}\frac{dx^i}{ds}-\sum_i g'_{ir}\frac{dx^i_1}{ds}=u_r-u'_r.$$

The left side of this equation is the value when $\epsilon^r=0$ of the differential coefficient with respect to ϵ^r of the total length of the varied line. The first term on the right is the limit when $\epsilon^r=0$ of the ratio to ϵ^r of the projection on the original line of the variation in the position of P, and the second term is the similar ratio for P'. The entire right side, therefore, is the value when $\epsilon^r=0$ of the differential coefficient with respect to ϵ^r of the length of the projection of the varied line on the original line. The equation therefore asserts that, for small values of ϵ^r, the length of the varied line and that of its projection on the original line remain equal.

A FUNCTIONAL EQUATION FROM THE THEORY OF THE RIEMANN $\zeta(s)$-FUNCTION.

By A. Arwin.

In the following brief paper a functional equality is deduced by means of which the well-known Riemann formula from the analytic theory of prime numbers may be obtained. The method which has been followed in this connection may be followed also, without any difficulty, in the case of more general ζ-functions.

In Acta Mathematica vol. 25, p. 166, Mellin establishes the formula

$$(1)\qquad \log(1+x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \frac{\pi}{\sin \pi s}\cdot\frac{x^s}{s}\,ds, \qquad (0 < a < 1),$$

from which he derives the result*

$$(2)\qquad \log \pi(x) = \frac{1}{2\pi i}\int_{a-i\infty}^{a+i\infty} \frac{\pi}{\sin \pi s}\, S(s)\,\frac{x^s}{s}\,ds$$

where

$$(3)\qquad \pi(x) = \prod_1^\infty\left[\left(1+\frac{x}{a_\nu}\right) R^{-\frac{x}{a_\nu}+\frac{1}{2}\left(\frac{x}{a_\nu}\right)^2-\dots+(-1)^p\frac{1}{p}\left(\frac{x}{a_\nu}\right)^p}\right]^{m_\nu}$$

and

$$S(s) = \sum_{(\nu)} \frac{m_\nu}{a_\nu^s}, \qquad (\varrho < a < p+1)$$

p being the rank of $\pi(x)$ and ϱ the exponent of convergence of $\sum m_\nu/a_\nu$.

From

$$(4)\qquad \pi(x) = \prod_{(n,m)}\left[\left(1+\frac{x}{p_n^m}\right) R^{-\frac{x}{p_n^m}}\right]^{\frac{\log p_n}{2\pi}}$$

we find

$$(5)\qquad S(s) = \frac{1}{2\pi}\sum_{(n,m)}\frac{\log p_n}{p_n^{ms}} = -\frac{1}{2\pi}\,\frac{\zeta'(s)}{\zeta(s)};$$

* Mellin, H., Acta Math., vol. 25, p. 168.

where p_n means the nth prime number in the series of primes. Then we have

$$(6)\qquad -2\pi\sum_{(n,m)}\left[\frac{\log p_n}{2\pi}\log\left(1+\frac{x}{p_n^m}\right)-\frac{\log p_n}{2\pi}\,\frac{x}{p_n^m}\right] = \frac{1}{2\pi i}\int\limits_{a-i\infty}^{a+i\infty}\frac{\pi}{\sin\pi s}\cdot\frac{\zeta'(s)}{\zeta(s)}\cdot\frac{x^s}{s}\,ds$$

where $1<a<2$, $x=|x|\,e^{i\theta}$, $-\pi<\theta<\pi$.

We now make a cut through the negative real axis and take the integral

$$(2')\qquad \frac{1}{2\pi i}\int\frac{\pi}{\sin\pi s}\cdot\frac{\zeta'(s)}{\zeta(s)}\cdot\frac{x^s}{s}\,ds$$

around the rectangle c_τ running from $a-i\tau$ to $a+i\tau$ to $-\sigma+i\tau$ to $-\sigma-i\tau$ back to $a-i\tau$. Since*

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right|<k_1\log|s|$$

for $s=\sigma+it$, $(\sigma=-3,-5,\ldots)$, we conclude that the integral from $-\sigma+it$ to $-\sigma-it$ vanishes for $s\to\infty$. Moreover†

$$\left|\frac{\zeta'(s)}{\zeta(s)}\right|<k_2\log^2|s|$$

at all points on all lines parallel to the real axis. Hence the integral along those sides of c_τ which are parallel to the real axis will also approach zero as $\tau\to\infty$. Hence we get

$$(7)\qquad -\sum_{(n,m)}\left[\log p_n\log\left(1+\frac{x}{p_n^m}\right)-\log p_n\frac{x}{p_n^m}\right]=\operatorname{Res}\left[\frac{\pi}{\sin\pi s}\cdot\frac{\zeta'(s)}{\zeta(s)}\cdot\frac{x^s}{s}\right]_{c_\infty},$$

the residue being evaluated for the region c_∞, i. e., for c_τ as $\sigma\to\infty$ $\tau\to\infty$.

* Landau, E., Handbuch der Lehre von der Verteilung der Primzahlen, vol. I, p. 336.
† Landau, E., ibid., p. 341.

In the first place

$$(8)\qquad \begin{aligned}\operatorname*{Res}_{s=0}\left[\frac{\pi}{\sin\pi s}\cdot\frac{\zeta'(s)}{\zeta(s)}\cdot\frac{x^s}{s}\right] &= \operatorname*{Res}_{s=0}\left[\frac{\zeta'(s)}{\zeta(s)}\cdot\frac{x^s}{s^2}\right]\\ &= \frac{\zeta'(0)}{\zeta(0)}\log x+\left(\frac{\zeta''(0)}{\zeta(0)}-\left(\frac{\zeta'(0)}{\zeta(0)}\right)^2\right) = \log 2\pi\log x + k_3.\end{aligned}$$

Furthermore, for $s = 1$, we have*

$$\frac{\zeta'(s)}{\zeta(s)} = -\frac{1}{s-1}+E+A_1(s-1)+\cdots,$$

$$\frac{\pi}{\sin\pi s} = -\frac{1}{s-1}+B_1(s-1)+\cdots;$$

hence

$$(9)\qquad \begin{aligned}&\operatorname*{Res}_{s=1}\left[\frac{\pi}{\sin\pi s}\cdot\frac{\zeta'(s)}{\zeta(s)}\cdot\frac{x^s}{s}\right]\\ &= \operatorname*{Res}_{s=1}\left[\frac{x^s}{s(s-1)^2}-\frac{Ex^s}{s(s-1)}\right] = x\log x - x(1+E).\end{aligned}$$

We turn next to consider the residue at the set of points $s = -1, -2, \ldots$, which are double poles for even values of s, and only simple poles for odd values of s. By use of the well-known relation†

$$\pi\zeta(s) = (2\pi)^s\sin\frac{\pi s}{2}\,\Gamma(1-s)\,\zeta(1-s),$$

it follows that

$$\frac{\zeta'(s)}{\zeta(s)} = \log 2\pi+\frac{\pi}{2}\cot\frac{\pi s}{2}-\frac{\Gamma'(1-s)}{\Gamma(1-s)}-\frac{\zeta'(1-s)}{\zeta(1-s)}.$$

Now,

$$(10)\qquad \operatorname*{Res}_{s=-1,\ldots}\left[\frac{\pi}{\sin\pi s}\log 2\pi\,\frac{x^s}{s}\right] = \log 2\pi\log\left(1+\frac{1}{x}\right),$$

---

* Landau, E., ibid., vol. 1, p. 165.

† Landau, E., ibid., vol. 1, p. 285.

and since*

$$\frac{\pi^2}{\sin^2 \pi z} = \sum_{-\infty}^{\infty} \frac{1}{(z+\nu)^2},$$

or

$$\frac{\pi^2}{4\left(\sin \frac{\pi s}{2}\right)^2} = \sum_{-\infty}^{\infty} \frac{1}{(s+2\nu)^2},$$

we may write

$$\frac{\pi}{2} \operatorname*{Res}_{s=-1,\ldots} \left| \frac{\pi x^s}{s} \cdot \frac{\cot \frac{\pi s}{2}}{\sin \pi s} \right|_{c_\infty} = \operatorname{Res} \sum_{1}^{\infty} \left[\frac{1}{(s+2\nu)^2} \cdot \frac{x^s}{s} \right]$$

$$= -\log x \sum_{1}^{\infty} \frac{x^{-2\nu}}{2\nu} - \sum_{1}^{\infty} \frac{x^{-2\nu}}{4\nu^2} = \frac{1}{2} \log x \log\left(1 - \frac{1}{x^2}\right) - \sum_{1}^{\infty} \frac{x^{-2\nu}}{4\nu^2}.$$

To simplify this expression, we start from the formula

$$-\sum_{1}^{\infty} \frac{x^{-2\nu}}{2\nu} = \frac{1}{2} \log\left(1 - \frac{1}{x^2}\right).$$

Multiplying through by $1/x$ and integrating from x to ∞, we find

$$-\sum_{1}^{\infty} \int_{x}^{\infty} \frac{x^{-(2\nu+1)}}{2\nu}\, dx = \frac{1}{2} \int_{x}^{\infty} \frac{dx}{x} \log\left(1 - \frac{1}{x^2}\right);$$

hence

$$-\sum_{1}^{\infty} \frac{x^{-2\nu}}{4\nu^2} = -\frac{1}{2} \log x \log\left(1 - \frac{1}{x^2}\right) - \int_{x}^{\infty} \frac{\log x\, dx}{x(x^2-1)}.$$

We may therefore write

$$(11) \qquad \frac{\pi^2}{2} \operatorname*{Res}_{s=-1,\ldots} \left| \frac{\cot \frac{\pi s}{2}}{\sin \pi s} \frac{x^s}{s} \right| = -\int_{x}^{\infty} \frac{\log x\, dx}{x(x^2-1)}.$$

Further, we have

$$-\pi \operatorname*{Res}_{s=-1,\ldots} \left[\frac{\Gamma'(1-s)}{\Gamma(1-s)} \frac{x^s}{s} \frac{1}{\sin \pi s} \right]_{c_\infty} = \sum_{1}^{\infty} (-1)^n \frac{\Gamma'(n+1)}{\Gamma(n+1)} \frac{x^{-n}}{n},$$

* Lindelöf, E., Calcul des résidues, p. 54.

and

$$-\pi \operatorname*{Res}_{s=-1,\ldots}\left[\frac{\zeta'(1-s)}{\zeta(1-s)}\,\frac{x^s}{s}\,\frac{1}{\sin\pi s}\right]=\sum_1^{\infty}(-1)^n\frac{\zeta'(n+1)}{\zeta(n+1)}\,\frac{x^{-n}}{n}.$$

But

$$\sum_1^{\infty}(-1)^n\frac{\zeta'(n+1)}{\zeta(n+1)}\,\frac{x^{-n}}{n}=\sum_{(n,m)}\frac{\log p_n}{p_n^m}\log\left(1+\frac{1}{x p_n^m}\right), \tag{12}$$

and

$$\operatorname*{Res}_{\zeta(s)=0}\left[\frac{\pi}{\sin\pi s}\,\frac{\zeta'(s)}{\zeta(s)}\,\frac{x^s}{s}\right]_{c_\infty}=\pi\sum_{(n)}\frac{x^{\varrho_n}}{\varrho_n\sin\pi\varrho_n}+\pi\sum_{(n)}\frac{x^{\bar\varrho_n}}{\bar\varrho_n\sin\pi\bar\varrho_n},$$

where $\varrho_n=\sigma_n+i\beta_n$ represent the complex roots of $\zeta(s)=0$.

From $x\Gamma(x)=\Gamma(x+1)$, we have

$$\frac{\Gamma'(m+1)}{\Gamma(m+1)}-\frac{\Gamma'(m)}{\Gamma(m)}=\frac{1}{m},$$

which, together with $\Gamma'(1)=-E$, $\Gamma(1)=1$, gives us the relation

$$\frac{\Gamma'(m+1)}{\Gamma(m+1)}=-E+\sum_1^m\frac{1}{n}.$$

Hence

$$\begin{aligned}\sum_1^{\infty}(-1)^n\frac{\Gamma'(n+1)}{\Gamma(n+1)}\,\frac{x^{-n}}{n}&=-E\sum_1^{\infty}(-1)^n\frac{x^{-n}}{n}+\sum_1^{\infty}(-1)^n\frac{x^{-n}}{n}\sum_1^n\frac{1}{m}\\ &=E\log\left(1+\frac{1}{x}\right)+\sum_1^{\infty}(-1)^n\frac{x^{-n}}{n^2}+\sum_1^{\infty}(-1)^n\frac{x^{-n}}{n}\sum_1^{n-1}\frac{1}{m}.\end{aligned} \tag{13}$$

We readily verify the relations

$$\sum_1^{\infty}(-1)^n\frac{x^{-n}}{n^2}=\log x\log\left(1+\frac{1}{x}\right)-\int_x^{\infty}\frac{\log x\,dx}{x(x+1)}, \tag{14'}$$

and

$$\sum_1^{\infty}(-1)^n\frac{x^{-n}}{n}\sum_1^{n-1}\frac{1}{m}=\frac{1}{2}\left[\log\left(1+\frac{1}{x}\right)\right]^2. \tag{14''}$$

Combining the results of equations (7) to (14), we arrive at the following relation*

* Wigert, S., "Sur la théorie de la fonction $\zeta(s)$ de Riemann", Arkiv för Mat., Astr. och Fysik., vol. 14, 1919. In this paper Wigert has deduced an equation which follows from relation (15). Circumstances have prevented earlier publication of this formula, which I obtained as early as 1917.

$$-\sum_{(n,m)}\left[\frac{\log p_n}{p_n^m}\log\left(1+\frac{1}{x p_n^m}\right)\right]-\sum_{(n,m)}\left[\log p_n \log\left(1+\frac{x}{p_n^m}\right)-\log p_n\frac{x}{p_n^m}\right]$$

$$(15)\quad = k_3 + x\log x - x(1+E) + E\log\left(1+\frac{1}{x}\right) + \log 2\pi\log(1+x)$$

$$+\log x\log\left(1+\frac{1}{x}\right)+\frac{1}{2}\left\{\log\left(1+\frac{1}{x}\right)\right\}^2-\int_x^\infty\frac{\log x\,dx}{x^2-1}+\pi\sum_{(\rho_n)}\frac{x^{\rho_n}}{\varrho_n\sin\pi\varrho_n}.$$

Now let x be arbitrary but not real and substitute $xe^{-\pi i}$ for x in (15); then we have

$$\sum_{(n)}\frac{x^{\rho_n}}{\varrho_n}\,\frac{R^{-\pi i\rho_n}}{R^{\pi i\rho_n}-R^{-\pi i\rho_n}} = -\sum_{(n)}\frac{x^{\rho_n}}{\varrho_n}\left(1+\sum_{m=1}^{\infty}R^{2m\pi i\rho_n}\right)$$

and

$$\sum_{(n)}\frac{x^{\bar\rho_n}}{\bar\varrho_n}\,\frac{R^{-\pi i\bar\rho_n}}{R^{\pi i\bar\rho_n}-R^{-\pi i\bar\rho_n}} = \sum_{(n)}\frac{x^{\bar\rho_n}}{\varrho_n}\left(\sum_1^{\infty}e^{-2m\pi i\bar\rho_n}\right);$$

so that

$$(16)\quad \begin{aligned}&\pi\sum_{(n)}\frac{x^{\rho_n}R^{-\pi i\rho_n}}{\varrho_n\sin\pi\varrho_n}+\pi\sum_{(n)}\frac{x^{\bar\rho_n}R^{-\pi i\bar\rho_n}}{\bar\varrho_n\sin\pi\bar\varrho_n}\\ &\qquad = -2\pi i\sum_{(n)}\frac{x^{\rho_n}}{\varrho_n}+4\pi J_m\left(\sum_{m,n}\frac{x^{\rho_n}}{\varrho_n}R^{2m\pi i\rho_n}\right).\end{aligned}$$

Recalling now the well-known expression

$$N(\tau) = \frac{\tau}{2\pi}\log\tau-\frac{\tau}{2\pi}(1+\log 2\pi)+O(\log\tau),$$

and*

$$s_n = 2\pi\frac{n}{\log n}(1+\varepsilon(n)),$$

we may prove that the last series on the right side of (16) is absolutely and uniformly convergent for all finite x, real as well as complex. In this manner we obtain from (15) the equation

* Backlund, R. J., Über die Nullstellen der Riemannschen Zetafunktion, Helsingfors 1916, p. 11.

$$-\sum_{(n,m)}\left[\frac{\log p_n}{p_n^m}\log\left(1-\frac{1}{x p_n^m}\right)\right]-\sum_{(n,m)}\left[\log p_n \log\left(1-\frac{x}{p_n^m}\right)+\log p_n \frac{x}{p_n^m}\right]$$

$$= k_3+(1+E)\log x - x\log x + x\pi i+\log 2\pi \log(1-x)$$

$$(15') \qquad + E\log\left(1-\frac{1}{x}\right)+\frac{1}{2}\left\{\log\left(1-\frac{1}{x}\right)\right\}^2+\log x \log\left(1-\frac{1}{x}\right)$$

$$-\pi i \log\left(1-\frac{1}{x}\right)+\int_x^\infty \frac{\log x\, dx}{x^2-1}-\pi i\int_x^\infty \frac{dx}{x^2-1}$$

$$-2\pi i\sum_{(n)}\frac{x^{\rho_n}}{\varrho_n}+4\pi J_m\left(\sum_{m,n}\frac{x^{\rho_n}}{\varrho_n}R^{2m\pi i\rho_n}\right).$$

The formulas (15) and (15′) have only been deduced for the case that $+\pi > \arg x > -\pi$, $2\pi > \arg x > 0$, but, since Landau has proved the series*

$$(17) \qquad \sum_{(n)}\frac{x^{\rho_n}}{\varrho_n}$$

to be uniformly convergent in all intervals not containing any point $x = 1$, p_n^m, $1/p_n^m$, we infer that in fact (15) and (15′) are valid for all x. Further we infer that the series $\sum_{(n)} x^{\rho_n}/\varrho_n$ converges uniformly for all complex x-values, and we find also that the series (17) in the singular points $x = p_n^m$, $1/p_n^m$ will tend to infinity in the same way as the finite series $\sum_{s_n \leq \tau} x^{\rho_n}/\varrho_n$ for $x = p_n^m$, $1/p_n^m$ and $\tau \to \infty$. Going from the upper part of the complex plane towards the positive real axis, which now represents the cut-line, we obtain, by means of the equation

$$\sum_{(n,m)}\left[\log p_n \log\left(1-\frac{x}{p_n^m}\right)+\log p_n\frac{x}{p_n^m}\right]=\sum_{p_n^m>x}\left[\log p_n\log\left(1-\frac{x}{p_n^m}\right)+\log p_n\frac{x}{p_n^m}\right]$$

$$+\sum_{p_n^m<x}\left[\log p_n \log\left(\frac{1}{p_n^m}-\frac{1}{x}\right)+\log p_n\frac{x}{p_n^m}\right]+\log x\sum_{p_n^m<x}\log p_n-\pi i\sum_{p_n^m<x}\log p_n.$$

* Landau, E., Math. Annalen, vol. 71, p. 563.

and, by use of the formula

$$\log 2\pi \log(1-x) = \log 2\pi \log\left(1-\frac{1}{x}\right) + \log 2\pi \log x - \pi i \log 2\pi,$$

the relation

$$(18) \quad \sum_{p_n^m \leqq x} \log p_n = x - \log\left(1-\frac{1}{x}\right) - \int_x^\infty \frac{dx}{x^2-1} - \log 2\pi - 2R\sum_{(n)} \frac{x^{\rho_n}}{\rho_n}.$$

Since

$$\int_x^\infty \frac{dx}{x^2-1} + \log\left(1-\frac{1}{x}\right) = \tfrac{1}{2}\log\frac{x+1}{x-1} + \log(x-1) - \log x = \tfrac{1}{2}\log\left(1-\frac{1}{x}\right),$$

we reach the well-known form*

$$\sum_{p_n^m \leqq x} \log p_n = x - \tfrac{1}{2}\log\left(1-\frac{1}{x^2}\right) - \log 2\pi - 2R\sum_{(n)} \frac{x^{\rho_n}}{\rho_n}.$$

It may be observed that we might have deduced (18) without referring to the above mentioned results of Landau, namely by modifying the manner of obtaining the limiting series which results if we approach the real axis. Substituting in (15) first $1/x$ and then $xe^{-\pi i}$ for x, we obtain, in the same way as before, the expression

$$(19) \quad \sum_{p_n^m \leqq x} \frac{\log p_n}{p_n^m} = \log x - \frac{1}{x} + \frac{1}{2}\log\frac{x+1}{x-1} - E + 2R\sum_{(n)} \frac{\left(\frac{1}{x}\right)^{\rho_n}}{\rho_n}.$$

The relation (19) can be verified directly, except as concerns the constant, by some formulas of transformation which have been previously published.†

In conclusion, we repeat that all these formulas may be deduced also for the case of more general ζ-functions.

LUND, SWEDEN,
November 1921.

* Landau, E., Handbuch etc., vol. 1, p. 353.

† Arwin, A., Nyt Tidskrift for Math., vol. 27, p. 81, formula (II).

THE GEOMETRY OF PATHS AND GENERAL RELATIVITY.

By Luther Pfahler Eisenhart.

Introduction. The geometry of paths, as developed by Professor Veblen and myself in a number of papers in volumes eight and nine of the Proceedings of the National Academy of Sciences, is a generalization of Euclidean geometry based on the assumptions that the space considered is an n-dimensional continuum in the sense of Analysis Situs and that in this space there exists a system of curves, called *paths*, such that, like the straight lines of Euclidean space, there is a unique path through any point in any direction. In the first part of this paper some of the results previously obtained are coördinated with certain new ones.

The coefficients $\Gamma^{i}_{\alpha\beta}$ in the equations (1.1) of the paths may be used to define the affine connection of the manifold, as a generalization of Levi-Civita's concept of infinitesimal parallelism in a Riemannian manifold. In this way the geometry of paths may be identified with the generalized geometries studied by Weyl* and Eddington.† These writers make the concept of parallelism fundamental in the development of their geometries, whereas in the geometry of paths the equations of the paths are fundamental.

In the second part of the paper we assume that the space-time continuum of physics is a four dimensional geometry of paths, elemental physical phenomena manifesting themselves in paths. It is generally accepted that the world-lines of light through a point form a quadratic cone. This leads us to adopt a particular form for the functions $\Gamma^{i}_{\alpha\beta}$ in the equations of paths, which we take as characterizing the geometry of paths of physics. This geometry includes, as a particular case, the geometry proposed by Weyl in order to obtain a metric involving both gravitation and electricity. Einstein‡ has said that "a theory of relativity in which the gravitational field and the electromagnetic field enter as an essential unity" is desirable and recently has proposed such a theory.§ His geometry also is included in the one now proposed, and it may be that the latter, because of its greater generality and adaptability,

---

* Space, Time and Matter, § 14.

† The Mathematical Theory of Relativity, Chapter 7. The statement on page 242 that the geometry of paths is identical with Weyl's geometry involving a metric is not correct, as is seen when a comparison is made of the results of §§ 2-4 of this paper with Eddington's §§ 91-94.

‡ The Meaning of Relativity, p. 108.

§ Sitz. Preuß. Akad. Wiss. 1923, p. 32.

will serve better as the basis for the mathematical formulation of the results of physical experiment.

In § 9 we consider the motion of a charged particle in an electromagnetic field and show how the results of experiment are in keeping with the above assumptions. In the closing section we consider the energy-momentum tensor of matter.

I. Geometry of paths.

1. **Equations of paths.** The paths are, by definition, the curves in a general manifold of n dimensions defined by the system of equations

$$\frac{d^2 x^i}{ds^2} + \Gamma^i_{\alpha\beta} \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} = 0 \qquad (i = 1, 2, \ldots, n), \tag{1.1}$$

where $\Gamma^i_{\alpha\beta}\ (= \Gamma^i_{\beta\alpha})$ are functions of the x's, and s is a fundamental parameter peculiar to each path whose significance will be pointed out. In (1.1) it is understood that α and β are summed from 1 to n in accordance with the usual convention; this convention will be followed throughout the paper.

If s for any path is replaced in (1.1) by a function of any other parameter, t, then for this path the equations become

$$\frac{\dfrac{d^2 x^i}{dt^2} + \Gamma^i_{\alpha\beta} \dfrac{dx^\alpha}{dt} \dfrac{dx^\beta}{dt}}{\dfrac{dx^i}{dt}} = -\frac{\dfrac{d^2 t}{ds^2}}{\left(\dfrac{dt}{ds}\right)^2}. \tag{1.2}$$

From these n equations follow the $n(n-1)/2$ equations

$$\frac{dx^j}{dt}\frac{d^2 x^i}{dt^2} - \frac{dx^i}{dt}\frac{d^2 x^j}{dt^2} + \left(\Gamma^i_{\alpha\beta}\frac{dx^j}{dt} - \Gamma^j_{\alpha\beta}\frac{dx^i}{dt}\right)\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt} = 0, \tag{1.3}$$

of which $n-1$ are independent.

Any one path may be defined by $n-1$ equations of the form $\varphi_i(x^1, \ldots, x^n) = 0$, from which it follows that all the coördinates are expressible in terms of any one as parameter. Thus any one coördinate, say x^1, may be used as parameter for all the paths, and so also may t, where $x^1 = f(t)$, the function f being arbitrary.

If we have a particular path, that is, a particular solution of (1.3), the left-hand members of (1.2) for $i = 1, \dots n$ are equal to one another in consequence of (1.3), and are reducible to a function of t, say $\varphi(t)$. Hence from (1.2) we obtain an equation of the form

$$\varphi(t) = \frac{\dfrac{d^2 t}{ds^2}}{\left(\dfrac{dt}{ds}\right)^2}.$$

If $t = f(s)$ is a solution of this equation, and the independent variable t is replaced by this function of s, equations (1.2) are reducible to the form (1.1) for this particular solution. Hence:

The equations (1.1) *of paths may be replaced by* (1.3) *in which t is a parameter the same for all the paths: conversely each solution of* (1.3) *leads to the determination of a parameter s in terms of which the differential equations of the corresponding curve are reducible to the form* (1.1).*

2. **The curvature tensor.** If we put $x^i = \varphi^i(x'^1, \dots, x'^n)$, thus introducing a new set of coördinates, equations (1.1) become

$$\frac{d^2 x'^i}{ds^2} + \Gamma'^i_{jk} \frac{dx'^j}{ds} \frac{dx'^k}{ds} = 0, \tag{2.1}$$

where

$$\frac{\partial^2 x^p}{\partial x'^i \partial x'^j} + \Gamma^p_{qr} \frac{\partial x^q}{\partial x'^i} \frac{\partial x^r}{\partial x'^j} = \Gamma'^t_{ij} \frac{\partial x^p}{\partial x'^t}. \tag{2.2}$$

Expressing the condition of integrability of these equations regarded as differential equations for determining the x's as functions of the x''s, we obtain

$$\frac{\partial x^q}{\partial x'^i} \frac{\partial x^r}{\partial x'^j} \frac{\partial x^s}{\partial x'^k} B^p_{qrs} = \frac{\partial x^p}{\partial x'^t} B'^t_{ijk}, \tag{2.3}$$

where

$$B^p_{qrs} = \frac{\partial \Gamma^p_{qs}}{\partial x^r} - \frac{\partial \Gamma^p_{qr}}{\partial x^s} + \Gamma^p_{\alpha r} \Gamma^\alpha_{qs} - \Gamma^p_{\alpha s} \Gamma^\alpha_{qr}. \tag{2.4}$$

* Mr. J. L. Synge, of the University of Toronto, made the suggestion, in a letter to Professor Veblen, that the equations of the paths be written in the form (1.3) in terms of a parameter other than s.

Since equation (2.3) may be written in the form

$$\frac{\partial x^q}{\partial x'^i}\,\frac{\partial x^r}{\partial x'^j}\,\frac{\partial x^s}{\partial x'^k}\,\frac{\partial x'^t}{\partial x^p}\,B^p_{qrs} = B'^t_{ijk}, \tag{2.5}$$

then B^p_{qrs}, defined by (2.4), are the components of a tensor, which is contravariant of the first order and covariant of the third order. It is known as the *curvature tensor.**

From (2.4) follows

$$B^p_{qrs} = -B^p_{qsr}, \tag{2.6}$$

that is, the tensor is skew-symmetrical in r and s.

If in (2.2) we replace Γ^p_{qr} by the Christoffel symbol of the second kind† $\begin{Bmatrix} p \\ qr \end{Bmatrix}$ formed with respect to the fundamental form of a Riemann space, we obtain the well-known equations for such a space. These equations are the basis of covariant differentiation in a Riemann space. It follows directly that the theory of covariant differentiation can be generalized to the geometry of paths by replacing the Christoffel symbols by the corresponding Γ's. Thus, if b_{ij} is a covariant tensor of the second order, its covariant derivative b_{ijk} is given by

$$b_{ijk} = \frac{\partial b_{ij}}{\partial x^k} - b_{aj}\,\Gamma^a_{ik} - b_{ia}\,\Gamma^a_{jk}. \tag{2.7}$$

3. **Fundamental tensors of the second order.** If in (2.4) we contract for p and s, we obtain

$$B_{qr} = B^p_{qrp} = \frac{\partial \Gamma^p_{pq}}{\partial x^r} - \frac{\partial \Gamma^p_{qr}}{\partial x^p} + \Gamma^p_{ar}\,\Gamma^a_{qp} - \Gamma^p_{pa}\,\Gamma^a_{qr}. \tag{3.1}$$

Since

$$B_{qr} = \frac{1}{2}(B_{qr} + B_{rq}) + \frac{1}{2}(B_{qr} - B_{rq}),$$

* Weyl (l. c. p. 118) and Eddington (p. 198) arrive at this tensor by means of the parallel displacement of a vector.

† The usual notation is $\begin{Bmatrix} qr \\ p \end{Bmatrix}$, but we use the form in the text because the location of the indices is more in keeping with the convention of summation.

if we put

$$B_{qr} = b_{qr} + \varphi_{qr}, \tag{3.2}$$

the symmetric and skew-symmetric parts of B_{qr} are given by

$$b_{qr} = \frac{1}{2}\left(\frac{\partial \Gamma^{\alpha}_{\alpha q}}{\partial x^{r}} + \frac{\partial \Gamma^{\alpha}_{\alpha r}}{\partial x^{q}}\right) - \frac{\partial \Gamma^{\alpha}_{qr}}{\partial x^{\alpha}} + \Gamma^{\beta}_{\alpha r}\,\Gamma^{\alpha}_{q\beta} - \Gamma^{\beta}_{\beta\alpha}\,\Gamma^{\alpha}_{qr} \tag{3.3}$$

and

$$\varphi_{qr} = \frac{1}{2}\left(\frac{\partial \Gamma^{\alpha}_{\alpha q}}{\partial x^{r}} - \frac{\partial \Gamma^{\alpha}_{\alpha r}}{\partial x^{q}}\right). \tag{3.4}$$

Thus b_{qr} and φ_{qr} are the components of a symmetric and a skew-symmetric tensor respectively of the second order, which are fundamental in the geometry of paths. Moreover, from (2.4) it follows that

$$B^{\alpha}_{\alpha rs} = -2\,\varphi_{rs}. \tag{3.5}$$

It is well-known in the theory of tensors that δ^j_i, where

$$\delta^j_i = 0 \quad \text{for } i \neq j; \quad = 1, \quad \text{for } i = j, \tag{3.6}$$

are the components of a mixed tensor of the second order. If then a_{ij} are the components of any covariant tensor of the second order, a contravariant tensor of components $a^{(ij)}$ is defined by

$$a_{ik}\, a^{(kj)} = \delta^j_i. \tag{3.7}$$

Thus $a^{(kj)}$ is the cofactor of a_{kj} in the determinant

$$a = |a_{ij}| \tag{3.8}$$

divided by a. In like manner if a^{ik} are the components of a contravariant tensor, the components $a_{(ik)}$ of a covariant tensor are defined by

$$a^{ik}\, a_{(kj)} = \delta^i_j. \tag{3.9}$$

In either case we shall speak of the tensor derived in this manner as the *associate* tensor and we shall indicate it by the parenthesis about the indices.

We introduce the notation*

$$
(3.10)\qquad
\begin{aligned}
\Lambda_{ij.k} &= \frac{1}{2}\left(\frac{\partial b_{ik}}{\partial x^j} + \frac{\partial b_{jk}}{\partial x^i} - \frac{\partial b_{ij}}{\partial x^k}\right),\\
\Lambda^k_{ij} &= b^{k\alpha}\,\Lambda_{ij.\alpha}.
\end{aligned}
$$

From the first of these we have

$$
(3.11)\qquad \frac{\partial b_{ij}}{\partial x^k} = \Lambda_{ik.j} + \Lambda_{kj.i}.
$$

Also we have

$$
(3.12)\qquad \Lambda^i_{ij} = \frac{\partial \log \sqrt{b}}{\partial x^j},
$$

where

$$
(3.13)\qquad b = |b_{ij}|.
$$

Since the covariant derivative b_{ijk} is a covariant tensor of the third order, two fundamental covariant vectors β_j and b_j are defined by

$$
(3.14)\qquad b^{ik}\,b_{ikj} = -2\beta_j, \qquad b^{(ik)}\,b_{ijk} = -b_j.
$$

If in the first of these equations we substitute the expression for b_{ijk} as given by (2.7), the resulting equation is reducible by means of (3.11), (3.12) and (3.7) to

$$
(3.15)\qquad \Gamma^i_{ij} = \frac{\partial \log \sqrt{b}}{\partial x^j} + \beta_j.
$$

From this result and (3.4) we obtain

$$
(3.16)\qquad q_{qr} = \frac{1}{2}\left(\frac{\partial \beta_q}{\partial x^r} - \frac{\partial \beta_r}{\partial x^q}\right).
$$

* As thus defined $\Lambda_{ij.k}$ and Λ^k_{ij} are the Christoffel symbols of the first and second kinds formed with respect to the tensor b_{ij}.

Hence we have:

The skew-symmetrical tensor φ_{qr} of a geometry of paths is the curl of a vector.

From the second of (3.14) and (2.7) we have, with the aid of (3.11), (3.12) and (3.15),

$$(3.18)\qquad b^{ik}\, b_{aj}\,(\Lambda^a_{ik} - \Gamma^a_{ik}) = \beta_j - b_j.$$

Consequently

$$(3.19)\qquad \Gamma^a_{ik} = \Lambda^a_{ik} + a_{ik}{}^{a}.$$

where $a_{ik}{}^{a}$ is a tensor of the third order, contravariant of the first order and covariant of the second.*

From (3.19) and (3.15) we have

$$(3.20)\qquad \beta_j = a_{aj}{}^{a}.$$

and from (3.18) and (3.19)

$$(3.21)\qquad b_j = a^{a}{}_{aj} + a_{aj}{}^{a}.$$

From (3.3) and (3.15) we have

$$(3.22)\quad b_{qr} = \frac{\partial^2 \log \sqrt{b}}{\partial x^q\,\partial x^r} + \frac{1}{2}\left(\frac{\partial \beta_q}{\partial x^r} + \frac{\partial \beta_r}{\partial x^q}\right) - \frac{\partial \Gamma^p_{qr}}{\partial x^p} + \Gamma^p_{ar}\Gamma^a_{qp} - \Gamma^p_{pa}\Gamma^a_{qr}.$$

If we write

$$(3.23)\qquad B_{qr} = \frac{\partial^2 \log \sqrt{b}}{\partial x^q\,\partial x^r} - \frac{\partial \Lambda^p_{qr}}{\partial x^p} + \Lambda^p_{ar}\Lambda^a_{qp} - \Lambda^a_{qr}\frac{\partial \log \sqrt{b}}{\partial x^a},$$

then B_{qr} is the contracted Riemann tensor formed with respect to the tensor b_{qr}. Substituting from (3.15) and (3.19) in (3.22) and making use of (3.23), we obtain

$$(3.21)\qquad b_{qr} = \frac{1}{2}(\beta_{qr} + \beta_{rq}) + \bar{B}_{qr} - (a_{qr}{}^{a})_a + a_{aq}{}^{\beta}\, a_{\beta r}{}^{a}.$$

* This result follows directly from (2.2) and similar equations obtained by replacing the Γ's by the Λ's; the latter are the corresponding equations for a Riemann space with b_{rs} for fundamental form. Subtracting corresponding equations, we obtain equations whose form reveals the tensor character of $a_{ik}{}^{a}$.

where

$$(3.22)\qquad (a_{qr}{}^{\alpha})_{\alpha} = \frac{\partial a_{qr}{}^{\alpha}}{\partial x^{\alpha}} + a_{qr}{}^{\beta}\,\Gamma^{\alpha}_{\alpha\beta} - a_{\beta r}{}^{\alpha}\,\Gamma^{\beta}_{\alpha q} - a_{q\beta}{}^{\alpha}\,\Gamma^{\beta}_{\alpha r},$$

that is, the sum for α of the covariant derivative of $a_{qr}{}^{\alpha}$ with respect to the tensor b_{ij}, and β_{qr} is the covariant derivative of β_q as defined in § 2.

4. **Invariant integrals.** If b'_{ij} denote the components of the tensor b_{ij} in a set of coordinates x', we have

$$(4.1)\qquad b' = |b'_{ij}| = |b_{ij}|\,J^2 = b\,J^2,$$

where J is the Jacobian

$$(4.2)\qquad J = \left|\frac{\partial x^i}{\partial x'^j}\right|.$$

From (4.1) it follows that the integral

$$(4.3)\qquad \int \sqrt{b}\,dx^1 \ldots dx^n$$

is an invariant integral for transformations of the x's.

Suppose now that we have any mixed tensor of the second order whose components in x' and in x are denoted by $c'^{\,i}_{j}$ and c^i_j. We have

$$(4.4)\qquad |c'^{\,i}_{j}| = \left|c^p_q \frac{\partial x^q}{\partial x'^j}\frac{\partial x'^i}{\partial x^p}\right| = |c^p_q|\,J\,.\,\bar{J},$$

where

$$\bar{J} = \left|\frac{\partial x'^i}{\partial x^j}\right|.$$

Since $J\bar{J} = 1$, we have from (4.4)

The determinant of the components of any mixed tensor of the second order is an invariant scalar.

If c_{ij} are the components of any covariant tensor of the second order, then

$$(4.5)\qquad c^i_j = a^{(ip)}\,c_{pj},$$

where $a^{(ip)}$ is the associate of a tensor a_{ij}, is a mixed tensor of the second order. From the above theorem we have $|c^i_j| = c$, where c is an invariant scalar. From (4.5) we have

$$c = |a^{(ip)}| \, |c_{pj}| .$$

Multiplying by $|a| = |a_{pq}|$, we have

$$c \, |a_{pq}| = |c_{pq}| .$$

Hence we have

*The determinants of the components of any two covariant tensors of the second order differ at most by a scalar factor.**

As an immediate consequence of this theorem it follows that the invariant integral as (4.3) formed with respect to any covariant tensor of the second order differs from any other by having a scalar in the integrand.

An indirect proof of the above theorem, so far as symmetric tensors go, is obtained from the following considerations. If we take any symmetric tensor of the second order other than b_{ij}, say a_{ij}, and proceed as in § 3, we obtain in place of (3.15) the result

$$\text{(4.6)} \qquad \Gamma^i_{ij} = \frac{\partial \log \sqrt{a}}{\partial x^j} + \alpha_j, \quad \varphi_{rs} = \frac{1}{2}\left(\frac{\partial \alpha_r}{\partial x^s} - \frac{\partial \alpha_s}{\partial x^r}\right),$$

where a and α_j are analogous to b and β_j. From (3.16) and the second of (4.6) it follows that α_j and β_j differ at most by a gradient, say $\dfrac{\partial \log \varrho}{\partial x^j}$. Hence from (3.15) and (4.6), we have $\sqrt{a} = c\varrho\sqrt{b}$, where c is a constant.†

5. Different forms of the equations of paths. We inquire under what conditions the paths are defined by a second system of equations of the form (1.1), say

$$\text{(5.1)} \qquad \frac{d^2 x^i}{ds^2} + \overline{\Gamma}^i_{\alpha\beta} \frac{dx^\alpha}{ds} \frac{dx^\beta}{ds} = 0.$$

* Clearly this theorem does not apply to a tensor whose determinant is zero, for instance a tensor which is the product of two vectors.

† In the Proceedings of the National Academy of Sciences, vol. 9 (1923), pp. 4–7, I showed that a necessary and sufficient condition that a geometry of paths admit an integral is that φ_{rs} be the curl of a vector. The above results show that this condition is always satisfied. Eddington, l. c., p. 219, obtains an equivalent result, but his statement on p. 216 that $\Gamma^i_{i\mu} \partial x^\mu / \partial x'^\alpha$ is a vector is not correct, as follows from (3.15), since b is not a scalar invariant. Cf. Bull. Amer. Math. Soc., vol. 28 (1922), p. 427.

Making use of the same general parameter t, as in (1.3), we have

$$(5.2)\quad \frac{dx^j}{dt}\frac{d^2x^i}{dt^2}-\frac{dx^i}{dt}\frac{d^2x^j}{dt^2}+\left(\bar{\Gamma}^i_{\alpha\beta}\frac{dx^j}{dt}-\bar{\Gamma}^j_{\alpha\beta}\frac{dx^i}{dt}\right)\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt}=0.$$

Subtracting these equations, we obtain

$$\left[(\bar{\Gamma}^i_{\alpha\beta}-\Gamma^i_{\alpha\beta})\frac{dx^j}{dt}-(\bar{\Gamma}^j_{\alpha\beta}-\Gamma^j_{\alpha\beta})\frac{dx^i}{dt}\right]\frac{dx^\alpha}{dt}\frac{dx^\beta}{dt}=0.$$

This equation must be satisfied identically, otherwise the solution of (1.1) would satisfy an equation of the first order. It is readily found that a necessary and sufficient condition is that

$$(5.3)\quad \bar{\Gamma}^i_{\alpha\beta}=\Gamma^i_{\alpha\beta}+\psi_\alpha\,\delta^i_\beta+\psi_\beta\,\delta^i_\alpha.$$

If we subtract equations (2.2) from similar equations involving $\bar{\Gamma}^p_{qr}$ and $\bar{\Gamma}'^t_{ij}$, we obtain

$$\bar{\Gamma}'^t_{ij}-\Gamma'^t_{ij}=(\bar{\Gamma}^p_{qr}-\Gamma^p_{qr})\frac{\partial x^q}{\partial x'^i}\frac{\partial x^r}{\partial x'^j}\frac{\partial x'^t}{\partial x^p}.$$

Substituting from (5.3), we find that ψ_α are the components of a covariant vector.

Consider now a particular path and equations (1.2) and the analogous equations in terms of the barred functions. Subtracting these equations and making use of (5.3), we obtain

$$2\psi_\alpha\frac{dx^\alpha}{dt}=\frac{\dfrac{d^2t}{ds^2}}{\left(\dfrac{dt}{ds}\right)^2}-\frac{\dfrac{d^2t}{d\bar{s}^2}}{\left(\dfrac{dt}{d\bar{s}}\right)^2}.$$

Since s is a function of $\bar{s}$ along this curve, we have

$$\frac{dt}{d\bar{s}}=\frac{dt}{ds}\frac{ds}{d\bar{s}},\qquad \frac{d^2t}{d\bar{s}^2}=\frac{d^2t}{ds^2}\left(\frac{ds}{d\bar{s}}\right)^2+\frac{dt}{ds}\frac{d^2s}{d\bar{s}^2}.$$

Substituting in the above, we obtain

$$(5.4)\qquad \frac{\dfrac{d^2 s}{d\bar{s}^2}}{\left(\dfrac{ds}{d\bar{s}}\right)^2} = -2\,\psi_\alpha \frac{dx^\alpha}{ds},$$

for the determination of $\bar{s}$ as a function of s.

Hence we have the theorem:

If ψ_α are the components of any covariant vector, the paths defined by (1.1) *are also defined by* (5.1) *with the $\bar{\Gamma}$'s given by* (5.3) *and $\bar{s}$ by* (5.4).

If we denote by $\bar{B}^p_{qrs}$ the function in the $\bar{\Gamma}$'s analogous to (2.4), we have

$$(5.5)\quad \bar{B}^p_{qrs} = B^p_{qrs} + \delta^p_q(\psi_{sr} - \psi_{rs}) + \delta^p_s(\psi_{qr} - \psi_q\,\psi_r) - \delta^p_r(\psi_{qs} - \psi_q\,\psi_s).$$

where ψ_{rs} is the covariant derivative of ψ_r. Contracting for p and s, we obtain

$$(5.6)\qquad \bar{B}_{qr} = B_{qr} + n\,\psi_{qr} - \psi_{rq} - (n-1)\,\psi_q\,\psi_r.$$

Also contracting for p and q, we have

$$(5.7)\qquad \bar{\varphi}_{rs} = \varphi_{rs} + \frac{1}{2}(n+1)(\psi_{rs} - \psi_{sr}).$$

Since

$$\psi_{rs} - \psi_{sr} = \frac{\partial \psi_r}{\partial x^s} - \frac{\partial \psi_s}{\partial x^r},$$

it follows from (3.16) that, if we take

$$(5.8)\qquad (n+1)\,\psi_r = -\beta_r + \frac{\partial \sigma}{\partial x^r}, \qquad (r = 1, \ldots, n),$$

where σ is any function of the x's, then $\bar{\varphi}_{rs} = 0$.

Hence:

*The functions Γ and the parameter s may be chosen so that the skew-symmetric tensor is zero.**

From (5.5) and (5.6) we have that the tensor P^p_{qrs}, defined by

$$
\begin{aligned}
P^p_{qrs} = B^p_{qrs} &+ \frac{1}{n+1}\,\delta^p_q\,(B_{sr} - B_{rs}) \\
&+ \frac{1}{(n-1)(n+1)}\left[\delta^p_r\,(n\,B_{qs} + B_{sq}) - \delta^p_s\,(n\,B_{qr} + B_{rq})\right]
\end{aligned}
\tag{5.9}
$$

is independent of the vector ψ_i. This tensor was suggested by Weyl, who called it the *projective* tensor. In fact each vector ψ_i leads to a different affine connection, and consequently the geometry of paths may be looked upon as a generalized projective geometry, as Weyl has pointed out.†

From (5.9) we obtain

$$P^p_{qrp} = 0, \qquad P^p_{prs} = 0.$$

Thus only zero tensors of the second order are obtained by contraction.

6. **Normal and path coördinates.** The paths through a point P_0 of coördinates x^i_0 are completely determined by x^i_0 and the values of dx^i/ds at P_0. If we put

$$\left(\frac{dx^i}{ds}\right)_0 = \xi^i, \tag{6.1}$$

and choose s for each path so that $s = 0$ at P_0, the solutions of (1.1) are of the form

$$x^i = x^i_0 + \xi^i s - \frac{1}{2}(\Gamma^i_{\alpha\beta})_0\, \xi^\alpha \xi^\beta s^2 + \cdots, \tag{6.2}$$

* These results were obtained by me in a former paper, Proc. Nat. Acad. Scien., vol. 8 (1922), pp. 233–238; in the opening section of the paper I made the unnecessary requirement that s be the same for all paths; however, the results were not conditioned by that requirement. Later there was received in this country a paper by Weyl, Göttingen Nachrichten, 1921, containing similar results and extensions. Also Veblen, Proc. Nat. Acad. Scien., vol. 8 (1922), p. 347, obtained these results by a different method.

† L. c.; also, Veblen, l. c.

where $(\Gamma^i_{\alpha\beta})_0$ is the value of $\Gamma^i_{\alpha\beta}$ at P_0, and the other coefficients are formed by differentiating (1.1) successively and making use of (6.1).

If now we put

$$x'^i = \xi^i s \tag{6.3}$$

in (6.2), we obtain

$$x^i = x^i_0 + x'^i - \frac{1}{2}(\Gamma^i_{\alpha\beta})_0 x'^\alpha x'^\beta + \dots . \tag{6.4}$$

If these equations be solved for the x''s in terms of the x's, we obtain the equations of transformation of the coördinates x^i into coördinates x'^i in terms of which the equations of the paths are given by (6.3). Thus in terms of the x''s the equations of the paths through a particular point assume the parametric form of the equations of straight lines in Euclidean space. These coördinates are a generalization for the geometry of paths of the coördinates introduced by Riemann; they were first introduced by Veblen* and were called *normal* coördinates.

When the x''s are normal coördinates, it follows from (2.1) that at P_0

$$(\Gamma'^i_{jk})_0 = 0, \tag{6.5}$$

for all values i, j and k.

If in (6.4) the terms on the right after the third are dropped, these equations define a transformation of coördinates x^i into x'^i for which (6.5) holds at P_0. These coordinates are a generalization of geodesic coordinates in Riemannian geometry; we call them *path* coordinates. Thus normal coordinates are a special type of path coordinates.†

Since (6.5) are satisfied at P_0 for the path coördinates, the components of the first covariant derivatives of any tensor reduce at P_0 to the ordinary derivatives of the components of the latter.

Suppose now that we have any symmetric tensor of the second order g_{ij} and consider the quadratic form

$$g_{ij}\, dx^i\, dx^j \tag{6.6}$$

* Proc. Nat. Acad. Scien., vol. 8 (1922), pp. 192–197; the reader is referred to this paper for a fuller consideration of normal coördinates.

† From (6.4) and (5.3) it follows that, if the equations of the paths are taken in different forms in accordance with the results of § 5, normal or path coördinates for one form of the equations are not normal or path for another form. Hence, in using these coördinates at any time we are assuming that the equations of the paths have a definite form.

the coordinates x^i being normal. By a linear transformation with constant coefficients, the form (6.6) is reducible to

$$dx^{1^2} + \cdots + dx^{n^2} \tag{6.7}$$

at any point. Moreover, from (6.3) it follows that the new coördinates are normal. Hereafter we shall use the term normal coördinates in this restricted sense.

II. General relativity.

7. Paths of light. In accordance with the general theory of relativity the velocity of light is the same constant relatively to a local inertial system. If the units are chosen so that this velocity is unity, then in the local system the propagation of light is defined by the Galilean form

$$-dx^{1^2} - dx^{2^2} - dx^{3^2} + dx^{4^2} = 0, \tag{7.1}$$

where x^4 is the local coördinate of time. In general coördinates this assumes the form

$$g_{ij}\, dx^i\, dx^j = 0. \tag{7.2}$$

We have seen in § 6 that a system of normal path coordinates can be adopted with respect to this form so that at the point under consideration equation (7.2) becomes (7.1). A similar result can be obtained if normal Riemannian coördinates are used.* In the former case the equation

$$x^i = \xi^i s, \tag{7.3}$$

where the ξ's are constants, define paths; in the latter geodesics with respect to the form (7.2). As the former is less restricted, we adopt it. Consequently, if a four dimensional geometry of paths is to be identified with the space-time continuum of relativity, the functions $\Gamma^i_{\alpha\beta}$ must be such that there exists a symmetric tensor g_{ij} so that there shall be paths, namely the *paths of light*, satisfying (1.1) and

$$g_{ij} \frac{dx^i}{ds} \frac{dx^j}{ds} = 0. \tag{7.4}$$

* Cf. Birkhoff, Relativity and Modern Physics, pp. 118–124.

Differentiating this equation with respect to s and making use of (1.1), we obtain

$$g_{ijk}\frac{dx^i}{ds}\frac{dx^j}{ds}\frac{dx^k}{ds}=0, \tag{7.5}$$

where g_{ijk} is the covariant derivative of g_{ij} as defined in § 2.

Equation (7.5) is satisfied identically, if

$$g_{ijk}+g_{jki}+g_{kij}=0. \tag{7.6}$$

This is the condition that equations (1.1) admit the quadratic integral

$$g_{ij}\frac{dx^i}{ds}\frac{dx^j}{ds}=c, \tag{7.7}$$

where c is an arbitrary constant (cf. § 8).*

The condition (7.5) is satisfied by each path of light, if

$$g_{ijk}=g_{ij}\,\mu_k+g_{jk}\,\nu_i+g_{ki}\,\nu_j, \tag{7.8}$$

where μ_i and ν_i are the components of two covariant vectors.†

We denote by $[ij,k]$ and $\left\{ {k \atop ij} \right\}$ the Christoffel symbols of the first and second kinds formed with respect to g_{ij}. Thus

$$[ij,k]=\frac{1}{2}\left(\frac{\partial g_{ik}}{\partial x^j}+\frac{\partial g_{kj}}{\partial x^i}-\frac{\partial g_{ij}}{\partial x^k}\right), \tag{7.9}$$

$$\left\{ {k \atop ij} \right\}=g^{k\alpha}[ij,\alpha].\ddagger \tag{7.10}$$

* For a study of this case the reader is referred to a paper to be published by Veblen and Thomas, Transactions Amer. Math. Soc., vol. 25 (1923).

† This would be true also if ν_j in the last term were replaced by another vector σ_j, but then g_{ijk} would not be symmetric in i and j; the condition (7.5) would also be satisfied for each path of light, if there were added to (7.8) a term λc_{ijk}, where λ is a scalar invariant and c_{ijk} the covariant derivative of a symmetric tensor c_{ij} satisfying (7.6).

‡ Hereafter we use $g^{k\alpha}$ for $g^{(k\alpha)}$ as defined in § 3; also we make use of $g^{k\alpha}$ to raise indices and $g_{k\alpha}$ to lower them in the conventional manner.

From (7.8) it follows that

$$(7.11) \qquad [ij, k] = g_{\alpha k}\, \Gamma^{\alpha}_{ij} + \frac{1}{2}\left[g_{ij}(2\nu_k - \mu_k) + g_{jk}\,\mu_i + g_{ki}\,\mu_j\right],$$

and by means of (7.10)

$$(7.12) \qquad \Gamma^{k}_{ij} = \left\{ {k \atop ij} \right\} - \frac{1}{2}\left[g_{ij}(2\nu^k - \mu^k) + \delta^k_j\,\mu_i + \delta^k_i\,\mu_j\right].^{*}$$

From this we have

$$(7.13) \qquad \Gamma^{i}_{ij} = \frac{\partial \log \sqrt{-g}}{\partial x^j} - (\nu_j + 2\mu_j).$$

If, in accordance with § 4, we define a scalar invariant r by

$$(7.14) \qquad \sqrt{h} = r\sqrt{-g},$$

we have from (3.15) and (7.13)

$$(7.15) \qquad \beta_j = -\frac{\partial \log r}{\partial x^j} - (\nu_j - 2\mu_j).$$

Substituting the expressions (7.8) in (1.1), we obtain for the equations of general paths

$$(7.16) \quad \frac{d^2 x^i}{ds^2} + \left\{ {i \atop \alpha\beta} \right\} \frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = \frac{1}{2}(2\nu^i - \mu^i)\, g_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} - \mu_\alpha \frac{dx^\alpha}{ds}\frac{dx^i}{ds},$$

and, in consequence of (7.4), for the equations of paths of light

$$(7.17) \qquad \frac{d^2 x^i}{ds^2} + \left\{ {i \atop \alpha\beta} \right\} \frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = \mu_\alpha \frac{dx^\alpha}{ds}\frac{dx^i}{ds}.$$

* If we take $\nu_i = 0$, this assumes the form adopted by Weyl, Space, Time and Matter, p. 125; if we take $\nu_j = -\mu_j$, we get the form recently given by Einstein, Sitz. Preuß. Akad. Wiss., 1923, p. 36.

When the expressions (7.12) for the Γ's are substituted in (2.4), we obtain

$$
\begin{aligned}
B^p_{qrs} = G^p_{qrs} &+ \frac{1}{2}\,\delta^p_q(\mu_{rs} - \mu_{sr}) + \frac{1}{2}\,\delta^p_r\left(\mu_{qs} - \frac{1}{2}\,\mu_q\,\mu_s\right) \\
&- \frac{1}{2}\,\delta^p_s(\mu_{qr} - \mu_q\,\mu_r) + \frac{1}{2}\,g_{qr}\Big[2\,\nu^p_s - \mu^p_s + \frac{1}{2}(2\,\nu^p - \mu^p)\,\mu_s \\
&+ \frac{1}{2}\,\delta^p_s\,\mu_a(2\,\nu^a - \mu^a)\Big] - \frac{1}{2}\,g_{qs}\Big[2\,\nu^p_r - \mu^p_r \\
&\qquad + \frac{1}{2}(2\,\nu^p - \mu^p)\,\mu_r + \frac{1}{2}\,\delta^p_r\,\mu_a(2\,\nu^a - \mu^a)\Big].
\end{aligned}
\tag{7.18}
$$

where G^p_{qrs} is the Riemann tensor formed with respect to g_{ij}.

By means of (7.8) we have

$$
g_{qa}\,\nu^a_r = (g_{qa}\,\nu^a)_r - \nu^a\,g_{qar} = \nu_{qr} - \nu_q\,\mu_r - \nu_q\,\nu_r - g_{qr}\,\nu_a\,\nu^a.
$$

Making use of this result and contracting (7.18) for p and s, we obtain (3.2), where now

$$
\begin{aligned}
b_{qr} = G_{qr} &- \frac{1}{2}(\mu_{qr} + \mu_{rq}) + \frac{1}{2}\,\mu_q\,\mu_r - \frac{1}{2}(\nu_{qr} + \nu_{rq}) + \nu_q\,\nu_r \\
&+ \frac{1}{2}\,g_{qr}\left[2\,\nu^a_a - \mu^a_a + (2\,\nu^a - \mu^a)(\nu_a + 2\,\mu_a)\right].
\end{aligned}
\tag{7.19}
$$

$$
\varphi_{qr} = -\frac{1}{2}\left[2\,(\mu_{qr} - \mu_{rq}) + \nu_{qr} - \nu_{rq}\right].
\tag{7.20}
$$

G_{qr} being the contracted Riemann tensor for g_{ij}. Equation (7.20) is in conformity with (3.16) and (7.15).

The physical significance of the vectors μ_i and ν_i is a question which naturally presents itself. If they are assumed to vanish in empty space, equations (7.16) and (7.17) are the equations of geodesics. This conforms to

the case of the motion of Mercury and the deflection of light by the sun, if we put

$$G_{ij} = \lambda g_{ij}, \tag{7.21}$$

where λ is zero, or a very small constant. This equation follows from (7.19), if we assume for the general case that

$$b_{ij} = \lambda g_{ij} \tag{7.22}$$

as Eddington does.* Whether this is an unnecessary restriction can be determined only, it would seem, by further applications of the equations.

It may be that all physical phenomena can be accounted for by a single vector, as Weyl and Einstein have attempted to do, in place of the above two vectors. As will be shown in § 8, this can be accomplished by singling out a particular form of the equations of the paths, or, in other words, by a suitable determination of the affine connection of the continuum.

8. **Riemannian form of the equations of paths.** If equation (7.16) be multiplied by u_i, where $u^i = \frac{dx^i}{ds}$, and summed for i, we obtain

$$u_i\left(\frac{du^i}{ds} + \begin{Bmatrix} i \\ \alpha\beta \end{Bmatrix} u^\alpha u^\beta\right) = \frac{1}{2}(2\nu^\alpha + \mu^\alpha)u_\alpha u^i u_i. \tag{8.1}$$

Since we have identically

$$u^i \frac{du_i}{ds} = u^i \frac{d}{ds}(u^j g_{ij}) = u_j \frac{du^j}{ds} + 2u^\alpha u^\beta u_i \begin{Bmatrix} i \\ \alpha\beta \end{Bmatrix}, \tag{8.2}$$

equations (8.1) can be put in the form

$$\frac{d}{ds}(u^i u_i) = (2\nu^\alpha + \mu^\alpha)u_\alpha u^i u_i. \tag{8.3}$$

Hence if $u^i u_i = 0$ at a point of a path, it is zero along the path, and consequently the path is a path of light.

* L. c., p. 220; cf. also, Einstein, l. c., p. 35.

If we write the equations of the paths in terms of $\bar{\Gamma}^i_{\alpha\beta}$ and $\bar{s}$, given by (5.3) and (5.4), the $\bar{\Gamma}$'s assume a form analogous to (7.12), if we put

$$(8.4) \qquad \bar{\mu}_i = \mu_i - 2\psi_i, \qquad \bar{\nu}_i = \nu_i - \psi_i.$$

From these results it follows that if we take

$$(8.5) \qquad \psi_i = \frac{1}{4}(2\nu_i + \mu_i),$$

then $2\nu^a + \mu^a = 0$, and the right-hand member of the equation similar to (8.3) vanishes. Dropping the bars, we find that the equations of the paths may be taken in such a form that for paths other than light we have

$$(8.6) \qquad g_{\alpha\beta}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = 1.$$

We say that in this case the form of the equations of paths is *Riemannian*. This is equivalent to taking* in (7.8)

$$(8.7) \qquad 2\nu_i = -\mu_i.$$

Now (7.12) becomes

$$(8.8) \qquad \Gamma^k_{ij} = \left\{ {k \atop ij} \right\} + \frac{1}{2}[2g_{ij}\mu^k - \delta^k_j\mu_i - \delta^k_i\mu_j].$$

The equations of paths other than light are

$$(8.9) \qquad \frac{d^2x^i}{ds^2} + \left\{ {i \atop \alpha\beta} \right\}\frac{dx^\alpha}{ds}\frac{dx^\beta}{ds} = -\mu^i + \mu^a u_a u^i.$$

and the paths of light are given by (7.17).

---

* If in (8.4) we take $\psi^i = \nu^i$, we get the Weyl form, and $\psi_i = \mu_i + \nu_i$ leads to the Einstein form. It should be remarked that if g_{ij} satisfies (7.6), it is possible to choose s so that (8.6) holds for all paths other than light. Solutions of these equations other than (7.8), with (8.7) holding, will be given in another paper.

From (7.13) we have

$$\Gamma^i_{ij} = \frac{\partial \log \sqrt{-g}}{\partial x^j} - \frac{3}{2}\mu_j. \tag{8.10}$$

The expression (7.19) for b_{qr} reduces to

$$\begin{aligned} b_{qr} &= G_{qr} - \frac{1}{4}(\mu_{qr} + \mu_{rq}) + \frac{3}{4}\mu_q\mu_r - g_{qr}\left(\mu^a_a + \frac{3}{2}\mu^a\mu_a\right) \\ &= G_{qr} - \frac{1}{4}[(\mu_q)_r + (\mu_r)_q] + \frac{1}{4}\mu_q\mu_r + \frac{1}{2}g_{qr}[\mu_a\mu^a - 2(\mu^a)_a]. \end{aligned} \tag{8.11}$$

where $(\mu_q)_r$ denotes* the Riemannian covariant derivative of μ_q with respect to the form g_{ij}.

From (8.11) we have

$$g^{qr} b_{qr} = G + \frac{9}{4}[\mu_a\mu^a - 2(\mu^a)_a]. \tag{8.12}$$

9. **Motion of an electron in an electromagnetic field.** Maxwell's equations of the electromagnetic field, which are the foundation of the electron theory of Lorentz, state the relations holding between the electric force (X, Y, Z), the magnetic force (L, M, N), the density ϱ of electric charge and the density of electric current $(\sigma_x, \sigma_y, \sigma_z)$, where

$$\sigma_x = \varrho\frac{dx}{dt}, \qquad \sigma_y = \varrho\frac{dy}{dt}, \qquad \sigma_z = \varrho\frac{dz}{dt}. \tag{9.1}$$

The Maxwell equations can be given a simple and symmetric form by the introduction of a covariant vector (k_1, k_2, k_3, k_4) of which the first three components with signs changed are the components of the magnetic vector potential in the classical theory and k_4 is the electric potential. If we take the velocity of light as unity, and write

$$x^1 = x, \quad x^2 = y, \quad x^3 = z, \quad x^4 = t, \tag{9.2}$$

and

$$F_{ij} = \frac{\partial k_i}{\partial x^j} - \frac{\partial k_j}{\partial x^i}, \tag{9.3}$$

* This notation will be used hereafter.

the components of the electric and magnetic forces are given by the table:

$$
\begin{array}{c|cccc}
F_{ij} \quad \rightarrow i & & & & \\
\hline
\downarrow j & 0 & -N & M & -X \\
 & N & 0 & -L & -Y \\
 & -M & L & 0 & -Z \\
 & X & Y & Z & 0.
\end{array}
\tag{9.4}
$$

If we introduce the further notation

$$\varrho = \varrho_0 \frac{dt}{ds} \tag{9.5}$$

and

$$J^i = \varrho_0 \frac{dx^i}{ds}. \tag{9.6}$$

the Maxwell equations are

$$\frac{\partial F_{ij}}{\partial x^k} + \frac{\partial F_{jk}}{\partial x^i} + \frac{\partial F_{ki}}{\partial x^j} = 0. \tag{9.7}$$

$$g^{i\alpha}\, g^{j\beta} \frac{\partial F_{\alpha\beta}}{\partial x^j} = J^i, \tag{9.8}$$

where g_{ij} and g^{ij} have the Galilean values

$$
\begin{aligned}
g_{11} = g_{22} = g_{33} = -1, \quad g_{44} = 1, \quad g_{ij} = 0 \qquad (i \neq j), \\
g^{11} = g^{22} = g^{33} = -1, \quad g^{44} = 1, \quad g^{ij} = 0 \qquad (i \neq j).
\end{aligned}
\tag{9.9}
$$

Equations (9.7) and (9.8) determine the electromagnetic field when the distribution of currents and charges are known. The laws which govern the distribution of currents and charges are not known. However, it will be shown that the world-line of a slightly accelerated electron is a path, as defined in § 1.

Consider a slightly accelerated electron of stationary mass m and charge e, moving with velocity v in the direction of the x^1-axis at a point. By means of the Lorentz transformation and the resulting relations connecting the com-

ponents of the electric and magnetic forces as measured in two coordinate systems. Einstein* obtained the following equations

$$
\begin{aligned}
& m\beta^3 \frac{d^2x^1}{dt^2} = eX, \qquad \beta = \frac{1}{\sqrt{1-v^2}}, \qquad v = \frac{dx^1}{dt}, \\
(9.10) \qquad & m\beta \frac{d^2x^2}{dt^2} = e(Y - vN), \qquad \frac{dx^2}{dt} = 0, \\
& m\beta \frac{d^2x^3}{dt^2} = e(Z + vM), \qquad \frac{dx^3}{dt} = 0.
\end{aligned}
$$

These equations have been found to agree, to a high degree of accuracy, with the results of experiments upon β-particles.

Equations (9.10) may be written

$$
(9.11) \qquad \frac{d}{dt}\left(m\beta \frac{dx^1}{dt}\right) = eX, \qquad \frac{d}{dt}\left(m\beta \frac{dx^2}{dt}\right) = e(Y - vN), \ldots,
$$

and to them may be added

$$
(9.12) \qquad \frac{d}{dt}(m\beta) = evX.
$$

If we put

$$
v^i = \frac{dx^i}{dt} \quad (i = 1, 2, 3), \qquad v^4 = \frac{dt}{dt},
$$

equations (9.11) and (9.12) are a particular case of

$$
(9.13) \qquad
\begin{aligned}
& \frac{d}{dt}(m\beta v^i) = eF_{ij}v^j, \qquad (i = 1, 2, 3), \\
& \frac{d}{dt}(m\beta v^4) = -eF_{4j}v^j,
\end{aligned}
$$

where now

$$
(9.14) \qquad \beta = \frac{dt}{\sqrt{dt^2 - dx^{1^2} - dx^{2^2} - dx^{3^2}}}.
$$

* Ann. d. Phys., vol. 17 (1905), § 10.

If we put $t = x^4$ and write

$$w^i = \frac{dx^i}{d\sigma} \qquad (i = 1, 2, 3, 4), \tag{9.15}$$

where σ is a general parameter, equations (9.13) can be written in the form

$$\frac{d}{d\sigma}\left(\frac{m\,w^i}{\sqrt{g_{\alpha\beta}\,w^\alpha\,w^\beta}}\right) = -e\,g^{i\alpha}F_{\alpha j}\,w^j, \tag{9.16}$$

the g's having the Galilean values (9.9).

When the equations are written in the form (9.16), it becomes clear that in replacing equations (9.10) by (9.11) and (9.12) we made the tacit assumption that the derivatives of the g's are zero. Hence, in transforming (9.16) into general coördinates, we understand that in (9.16) the coördinates are geodesic. In effecting this transformation we make use of equations obtained from (2.2) by replacing Γ^i_{jk} by the Christoffel symbol $\left\{ {i \atop jk} \right\}$ formed with respect to the g's. This gives

$$\begin{aligned} \frac{m}{\sqrt{g_{\alpha\beta}\,w^\alpha\,w^\beta}}\left(\frac{dw^i}{d\sigma} + w^\alpha w^\beta \left\{ {i \atop \alpha\beta} \right\}\right) + m\,w^i \frac{d}{d\sigma}\frac{1}{\sqrt{g_{\alpha\beta}\,w^\alpha\,w^\beta}} \\ = -e\,g^{i\alpha}F_{\alpha j}\,w^j, \end{aligned} \tag{9.17}$$

where w^i is defined by (9.15), the x's being any coördinates whatever. Now

$$\frac{d}{d\sigma}\frac{1}{\sqrt{g_{\alpha\beta}\,w^\alpha\,w^\beta}} = -\frac{g_{jk}\,w^j}{(g_{\alpha\beta}\,w^\alpha\,w^\beta)^{\frac{3}{2}}}\left(\frac{dw^k}{d\sigma} + \left\{ {k \atop \alpha\beta} \right\} w^\alpha\,w^\beta\right).$$

Since F_{ij} is skew-symmetric, if we put

$$\frac{m}{\sqrt{g_{\alpha\beta}\,w^\alpha\,w^\beta}}\left(\frac{dw^i}{d\sigma} + \left\{ {i \atop \alpha\beta} \right\} w^\alpha\,w^\beta\right) + e\,g^{i\alpha}F_{\alpha j}\,w^j = A^i, \tag{9.18}$$

equation (9.17) reduces to

$$A^i\,w_\beta\,w^\beta = w^i\,w_\beta\,A^\beta.$$

If now we define a parameter s for the track of the particle by

$$\frac{d^2 s}{d\sigma^2} = \frac{w_\beta A^\beta}{m\sqrt{u^\beta w_\beta}} \frac{ds}{d\sigma},$$

and write

$$u^i = \frac{dx^i}{ds} \qquad (i = 1, \ldots, 4), \tag{9.19}$$

equations (9.18) become

$$\frac{m}{\sqrt{g_{\alpha\beta} u^\alpha u^\beta}} \left(\frac{du^i}{ds} + \left\{ {i \atop \alpha\beta} \right\} u^\alpha u^\beta\right) + e g^{i\alpha} F_{\alpha j} u^j = 0.$$

When this equation is multiplied by u_i and summed for i, we obtain

$$u_i \left(\frac{du^i}{ds} + \left\{ {i \atop \alpha\beta} \right\} u^\alpha u^\beta\right) = 0,$$

which, because of the identity (8.2), is reducible to $\frac{d}{ds}(u_i u^i) = 0$, and consequently s can be chosen so that (8.6) holds. Hence the above equation becomes

$$m \left(\frac{du^i}{ds} + \left\{ {i \atop \alpha\beta} \right\} u^\alpha u^\beta\right) + e F^{ij} u_j = 0. \tag{9.20}$$

In order that (9.20) be of the form (8.9), it is necessary and sufficient that

$$\mu^i = \frac{e}{m} F^{ij} u_j + \lambda u^i, \tag{9.21}$$

where λ is indeterminate. From (9.21) follows

$$\mu^i u_i = \mu_i u^i = \lambda. \tag{9.22}$$

If we take $\lambda = 0$, the vector μ_i is proportional to the ponderomotive force, acting on the particle. From (8.10), (3.15) and (3.16) it follows that the skew-symmetric tensor φ_{qr} appearing in the field is thus proportional to the

curl of the ponderomotive force vector and not of the electromagnetic potential vector, as has been assumed arbitrarily by Weyl, Eddington and Einstein.*

10. **Energy-momentum tensor of matter.** We write the energy-momentum tensor of matter in the form

$$T^{ij} = \sigma_0 \frac{dx^i}{ds}\frac{dx^j}{ds} = \sigma_0\, u^i u^j, \tag{10.1}$$

where σ_0 is the proper-density of matter at a place as estimated with reference to a system of coordinates moving with the matter, if we interpret ds to be the proper-time interval. In fact, if we put $u^i = 0$ ($i = 1, 2, 3$), $u^4 = 1$, we have $T^{44} = \sigma_0$. This interpretation of ds does not make it necessary that (8.6) shall hold: for from (8.3) we have, by choosing a suitable constant multiplier of s,

$$g_{\alpha\beta}\, u^\alpha u^\beta = e^{\int (2\nu_\alpha - \mu_\alpha)\, dx^\alpha}.$$

Hence it is only necessary that in the moving coordinate system $\nu_4 = \mu_4 = 0$.

From (10.1) we have

$$T^{ij}_{\ \ ,i} = |\sigma_0\, u^i|_i + \sigma_0\, u^i u^j_{,i}. \tag{10.2}$$

On the assumption that the world-lines are paths, the second term of the right-hand member vanishes. If we make the further assumption that

$$T^{ij}_{\ \ ,i} = 0 \tag{10.3}$$

everywhere, then from (10.2) we have

$$|\sigma_0\, u^i|_i = (\sigma_0\, u^i)_{,i} - 2\,\sigma_0\, u^i\,(\nu_i + 2\,\mu_i) = 0. \tag{10.4}$$

where the first term on the right is the sum for i of the Riemann covariant derivatives of $\sigma_0\, u^i$. This is a generalization of the equation of continuity. If the vector of components $\nu_i + 2\,\mu_i$ vanishes, that is, the vector whose curl is $2\,\varphi_{ij}$, by (3.16) and (7.15), this equation reduces to the equation of continuity.

* Weyl, Space, Time and Matter, p. 283; Eddington, l. c., pp. 201, 223; Einstein, Sitz. Preuß. Akad. Wiss., 1923, p. 35. See also a note by the author, Proc. Nat. Acad. Scien., vol. 9 (1923), pp. 175–178.

previously used in general relativity. By applying the rule of covariant differentiation, and making use of (7.15) we can put (10.3) in the form

$$(10.5)\quad \frac{\partial}{\partial x^j}(\sqrt{-g}\,T^{ij}) + \sqrt{-g}\,T^{ia}\,\Gamma^j_{ai} - \sqrt{-g}\,T^{aj}\,(\nu_j + 2\mu_j) = 0.$$

If we consider a particle in empty space, then ν_i and μ_i vanish except within the particle, and equations (10.5) assume the form

$$(10.6)\quad \frac{\partial}{\partial x^j}(\sqrt{-g}\,T^{ij}) + \sqrt{-g}\,T^{ia}\left\{ {j \atop ai} \right\} = 0.$$

By means of these equations and on the assumption that the particle is the nucleus of a symmetrical field, Eddington* has shown that the world-line of a particle is a geodesic, which is in agreement with the results of planetary motion. It may be too much to expect that (10.3), or its equivalent (10.5), describes the situation within the particle.

* L. c., pp. 126, 127.

PRINCETON UNIVERSITY.

INDEX

SEPTEMBER, 1922

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

PUBLISHED BY THE

PRINCETON UNIVERSITY PRESS

SECOND SERIES, VOL. 24, NO. 1

LANCASTER, PA., AND PRINCETON, N. J.

1923

According to an agreement between the Mathematical Association of America and the editors of the Annals of Mathematics, the Association contributes to the support of the Annals, and the Annals is supplied to individual members of the Association at one half of the regular price. In consequence of this agreement the volume of the Annals was increased by 100 pages, which are devoted to expository and historical articles in so far as suitable articles of this class are obtainable. Thus far the editors have not received enough such articles to fill the space available, and therefore wish to call the attention of authors to this lack and to the fact that, as long as the shortage continues, expository or historical articles of sufficient merit will receive prompt publication.

A number of the expository articles which have already been published are available in separate form and are listed for sale on the inside of the back cover of this number of the Annals. The regular subscription price of the Annals is $3.00 a volume.

Entered at the postoffice at Lancaster, Pa., as second-class matter under the Act of March 3, 1879.

CONTENTS

ANNALS OF MATHEMATICS

Published in September, December, March and June at Lancaster, Pa., and Princeton, N. J. Subscription price, \$3 a volume (four numbers) in advance. Single copies \$1.00. Subscriptions, orders for back numbers, and changes of address should be sent to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to The Annals of Mathematics, P. O. Box 53, Princeton, New Jersey. Manuscripts should be typewritten, with the exception of formulæ, and must be in final form with all references filled in.

Authors receive gratis 25 reprints of each article, postage prepaid. Additional copies will be furnished at cost.

Printed by Lütcke & Wulff, Hamburg, Germany

DECEMBER, 1922

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

EDITED BY

ORMOND STONE J. W. ALEXANDER
L. P. EISENHART T. H. GRONWALL
OSWALD VEBLEN J. H. M. WEDDERBURN

WITH THE COÖPERATION OF

A. A. BENNETT H. BLUMBERG
G. A. PFEIFFER J. K. WHITTEMORE

PUBLISHED BY THE

PRINCETON UNIVERSITY PRESS

SECOND SERIES, VOL. 24, NO. 2

LANCASTER, PA., AND PRINCETON, N. J.

1923

According to an agreement between the Mathematical Association of America and the editors of the Annals of Mathematics, the Association contributes to the support of the Annals, and the Annals is supplied at individual members of the Association at one half of the regular price. In consequence of this agreement the volume of the Annals was increased by 100 pages, which are devoted to expository and historical articles in so far as suitable articles of this class are obtainable. Thus far the editors have not received enough such articles to fill the space available, and therefore wish to call the attention of authors to this lack and to the fact that, as long as the shortage continues, expository or historical articles of sufficient merit will receive prompt publication.

A number of the expository articles which have already been published are available in separate form and are listed for sale on the inside of the back cover of this number of the Annals. The regular subscription price of the Annals is $3.00 a volume.

Entered at the postoffice at Lancaster, Pa., as second-class matter under the Act of March 3, 1879.

Copies of the following memoirs can be obtained by addressing THE ANNALS OF MATHEMATICS, Princeton, N. J.:

An elementary exposition of theory of the gamma function. By J. L. W. V. JENSEN. Authorized translation with additional notes by T. H. GRONWALL. 43 pages. Price 50 cents.

The gamma function in the integral calculus. By T. H. GRONWALL. 89 pages. Price 90 cents.

Fermat's last theorem and the origin and nature of the theory of algebraic numbers. By L. E. DICKSON. 27 pages. Price 35 cents.

Factorization of analytic functions of several variables. By W. F. OSGOOD. 19 pages. Price 25 cents.

Investigation of a class of fundamental inequalities in the theory of analytic functions. By J. L. W. V. JENSEN. Authorized translation from the Danish by T. H. GRONWALL. 29 pages. Price 40 cents.

CONTENTS

ANNALS OF MATHEMATICS

Published in September, December, March and June at Lancaster, Pa., and Princeton, N. J. Subscription price, $3 a volume (four numbers) in advance. Single copies $1.00. Subscriptions, orders for back numbers, and changes of address should be sent to the PRINCETON UNIVERSITY PRESS, PRINCETON, NEW JERSEY.

Manuscripts and all editorial correspondence should be addressed to THE ANNALS OF MATHEMATICS, P. O. Box 53, PRINCETON, NEW JERSEY. Manuscripts should be typewritten, with the exception of formulæ, and must be in final form with all references filled in.

Authors receive gratis 25 reprints of each article, postage prepaid. Additional copies will be furnished at cost.

PRINTED BY LÜTCKE & WULFF, HAMBURG, GERMANY

MARCH, 1923

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

PUBLISHED BY THE

PRINCETON UNIVERSITY PRESS

SECOND SERIES, VOL. 24, NO. 3

LANCASTER, PA., AND PRINCETON, N. J.

1923

According to an agreement between the Mathematical Association of America and the editors of the Annals of Mathematics, the Association contributes to the support of the Annals, and the Annals is supplied at individual members of the Association at one half of the regular price. In consequence of this agreement the volume of the Annals was increased by 100 pages, which are devoted to expository and historical articles in so far as suitable articles of this class are obtainable. Thus far the editors have not received enough such articles to fill the space available, and therefore wish to call the attention of authors to this lack and to the fact that, as long as the shortage continues, expository or historical articles of sufficient merit will receive prompt publication.

A number of the expository articles which have already been published are available in separate form and are listed for sale on the inside of the back cover of this number of the Annals. The regular subscription price of the Annals is $3.00 a volume.

Entered at the postoffice at Lancaster, Pa., as second-class matter under the Act of March 3, 1879.

CONTENTS

ANNALS OF MATHEMATICS

Published in September, December, March and June at Lancaster, Pa., and Princeton, N. J. Subscription price, $3 a volume (four numbers) in advance. Single copies $1.00. Subscriptions, orders for back numbers, and changes of address should be sent to the Princeton University Press, Princeton, New Jersey.

Manuscripts and all editorial correspondence should be addressed to The Annals of Mathematics, P. O. Box 53, Princeton, New Jersey. Manuscripts should be typewritten, with the exception of formulæ, and must be in final form with all references filled in.

Authors receive gratis 25 reprints of each article, postage prepaid. Additional copies will be furnished at cost.

Printed by Lütcke & Wulff, Hamburg, Germany

JUNE, 1923

ANNALS OF MATHEMATICS

(FOUNDED BY ORMOND STONE)

PUBLISHED BY THE

PRINCETON UNIVERSITY PRESS

SECOND SERIES, VOL. 24, NO. 4

LANCASTER, PA., AND PRINCETON, N. J.

1923

According to an agreement between the Mathematical Association of America and the editors of the Annals of Mathematics, the Association contributes to the support of the Annals, and the Annals is supplied at individual members of the Association at one half of the regular price. In consequence of this agreement the volume of the Annals was increased by 100 pages, which are devoted to expository and historical articles in so far as suitable articles of this class are obtainable. Thus far the editors have not received enough such articles to fill the space available, and therefore wish to call the attention of authors to this lack and to the fact that, as long as the shortage continues, expository or historical articles of sufficient merit will receive prompt publication.

A number of the expository articles which have already been published are available in separate form and are listed for sale on the inside of the back cover of this number of the Annals. The regular subscription price of the Annals is $3.00 a volume.

Entered at the postoffice at Lancaster, Pa., as second-class matter under the Act of March 3, 1879.

Copies of the following memoirs can be obtained by addressing THE PRINCETON UNIVERSITY PRESS, Princeton, N. J.:

An elementary exposition of theory of the gamma function. By J. L. W. V. JENSEN. Authorized translation with additional notes by T. H. GRONWALL. 43 pages. Price 50 cents.

The gamma function in the integral calculus. By T. H. GRONWALL. 89 pages. Price 90 cents.

Fermat's last theorem and the origin and nature of the theory of algebraic numbers. By L. E. DICKSON. 27 pages. Price 35 cents.

Factorization of analytic functions of several variables. By W. F. OSGOOD. 19 pages. Price 25 cents.

Investigation of a class of fundamental inequalities in the theory of analytic functions. By J. L. W. V. JENSEN. Authorized translation from the Danish by T. H. GRONWALL. 29 pages. Price 40 cents.

CONTENTS

ANNALS OF MATHEMATICS

Published in September, December, March and June at Lancaster, Pa., and Princeton, N. J. Subscription price, $3 a volume (four numbers) in advance. Single copies $1.00. Subscriptions, orders for back numbers, and changes of address should be sent to the PRINCETON UNIVERSITY PRESS, PRINCETON, NEW JERSEY.

Manuscripts and all editorial correspondence should be addressed to THE ANNALS OF MATHEMATICS, P. O. Box 53, PRINCETON, NEW JERSEY. Manuscripts should be typewritten, with the exception of formulæ, and must be in final form with all references filled in.

Authors receive gratis 25 reprints of each article, postage prepaid. Additional copies will be furnished at cost.

PRINTED BY LÜTCKE & WULFF, HAMBURG, GERMANY